OM nucleic - nucleic search, using sw model

```
                  GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.
```

Title: US-09-882-509-3
Perfect score: 1245
Sequence: 1 atgtgctgacctgagtggct.........ataccctaacgacaaataa 1245

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 294586 seqs, 226495651 residues

Total number of hits satisfying chosen parameters: 5883172

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Run on: May 6, 2004, 04:55:28 ; Search time 1265 Seconds
                                         (without alignments)
                4458.371 Million cell updates/sec

Database :   Published Applications NA:*
                1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
               10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
               11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
               12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
               13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
               14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
               15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
               16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
               17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
               18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
               19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

### SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1224.2 | 98.3 | 1245 | 10 | US-09-940-235-1 | Sequence 1, Appli |
| 2 | 1222.6 | 98.2 | 1377 | 10 | US-09-940-235-5 | Sequence 5, Appli |
| 3 | 1222.6 | 98.2 | 1782 | 10 | US-09-940-235-11 | Sequence 11, Appl |
| 4 | 1209.8 | 97.2 | 1327 | 10 | US-09-940-235-6 | Sequence 6, Appli |
| 5 | 1127.8 | 90.6 | 1661 | 10 | US-09-940-235-10 | Sequence 10, Appl |
| 6 | 1124.6 | 90.3 | 2096 | 10 | US-09-940-235-12 | Sequence 12, Appl |
| 7 | 1111.8 | 89.3 | 1541 | 10 | US-09-940-235-9 | Sequence 9, Appli |
| c 8 | 43.4 | 3.5 | 556 | 16 | US-10-369-493-29120 | Sequence 29120, A |
| c 9 | 42 | 3.4 | 5504 | 13 | US-10-221-613-270 | Sequence 270, App |
| c 10 | 42 | 3.4 | 6071 | 15 | US-10-311-455-340 | Sequence 340, App |
| c 11 | 42 | 3.4 | 6071 | 15 | US-10-240-485-36 | Sequence 36, Appl |
| c 12 | 42 | 3.4 | 3673778 | 15 | US-10-312-841-1 | Sequence 1, Appli |
| c 13 | 41.8 | 3.4 | 6332 | 15 | US-10-311-455-658 | Sequence 658, App |
| c 14 | 41.8 | 3.4 | 6392 | 15 | US-10-240-485-60 | Sequence 60, Appl |
| c 15 | 41.6 | 3.3 | 6283 | 15 | US-10-311-455-61 | Sequence 61, Appl |
| c 16 | 40.8 | 3.3 | 1794 | 13 | US-10-282-122A-40523 | Sequence 40523, A |
| c 17 | 40.8 | 3.2 | 9504 | 15 | US-10-240-453-281 | Sequence 281, App |
| c 18 | 40.4 | 3.2 | 45 | 10 | US-09-940-235-27 | Sequence 27, Appl |
| c 19 | 40.4 | 3.2 | 9084 | 15 | US-10-311-455-1578 | Sequence 1578, Ap |
| c 20 | 40.2 | 3.2 | 53 | 10 | US-09-940-235-14 | Sequence 14, Appl |
| 21 | 40.2 | 3.2 | 2000 | 9 | US-09-938-842A-3923 | Sequence 3923, Ap |
| 22 | 40.2 | 3.2 | 2000 | 11 | US-09-938-842A-3923 | Sequence 3923, Ap |
| 23 | 40.2 | 3.2 | 6407 | 15 | US-10-311-455-916 | Sequence 916, App |
| 24 | 40.2 | 3.2 | 8234 | 15 | US-10-311-455-1622 | Sequence 1622, Ap |
| 25 | 40.2 | 3.2 | 3673778 | 15 | US-10-312-841-2 | Sequence 2, Appli |
| c 26 | 39.6 | 3.2 | 13326 | 15 | US-10-311-455-1686 | Sequence 1686, Ap |
| c 27 | 39.4 | 3.2 | 19734 | 15 | US-10-311-455-1906 | Sequence 1906, Ap |
| c 28 | 39.4 | 3.2 | 580073 | 15 | US-10-205-220-1 | Sequence 1, Appli |
| c 29 | 38.8 | 3.1 | 5690 | 13 | US-10-240-454-26 | Sequence 26, Appl |
| c 30 | 38.8 | 3.1 | 5690 | 15 | US-10-311-455-1298 | Sequence 1298, Ap |
| c 31 | 38.8 | 3.1 | 5690 | 17 | US-10-257-166-110 | Sequence 110, App |
| c 32 | 38.8 | 3.1 | 16228 | 13 | US-10-221-613-387 | Sequence 387, App |
| c 33 | 38.6 | 3.1 | 640681 | 9 | US-09-790-988-1 | Sequence 1, Appli |
| c 34 | 38.4 | 3.1 | 1079 | 9 | US-09-070-927A-674 | Sequence 674, App |
| c 35 | 38.4 | 3.1 | 1890 | 13 | US-10-282-122A-21111 | Sequence 21111, A |
| c 36 | 38.2 | 3.1 | 609 | 10 | US-Q9-769-736-49 | Sequence 49, Appl |
| c 37 | 38.2 | 3.1 | 2232 | 15 | US-10-087-464-45 | Sequence 45, Appl |
| c 38 | 38 | 3.1 | 5514 | 15 | US-10-311-455-103 | Sequence 103, App |
| c 39 | 38 | 3.1 | 14095 | 15 | US-10-311-455-450 | Sequence 450, App |
| c 40 | 37.8 | 3.0 | 6104 | 13 | US-10-221-714A-17 | Sequence 17, Appl |
| c 41 | 37.8 | 3.0 | 6104 | 15 | US-10-311-455-269 | Sequence 269, App |
| c 42 | 37.8 | 3.0 | 8244 | 13 | US-10-221-714A-117 | Sequence 117, App |
| c 43 | 37.8 | 3.0 | 8244 | 15 | US-10-311-455-965 | Sequence 965, App |
| c 44 | 37.4 | 3.0 | 588 | 13 | US-10-027-632-265227 | Sequence 265227, |
| c 45 | 37.4 | 3.0 | 588 | 16 | US-10-027-632-265227 | Sequence 265227, |

### ALIGNMENTS

```
RESULT 1
US-09-940-235-1
; Sequence 1, Application US/09940235
; Publication_NO US2003005992lA1
; GENERAL INFORMATION:
;  APPLICANT: Kumar, Rajesh
;  APPLICANT: Sahni, Girish
;  APPLICANT: Roy, Chait
;  APPLICANT: Rajagopal, Kammara
;  APPLICANT: Nihalani, Deepak
;  APPLICANT: Sundaram, Vasudha
;  APPLICANT: Yadav, Mahavir
;  TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
;  TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
;  TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
;  TITLE OF INVENTION: PROTEIN
;  FILE REFERENCE: 07064-009002
;  CURRENT APPLICATION NUMBER: US/09/940,235
;  CURRENT FILING DATE: 2002-04-09
;  PRIOR APPLICATION NUMBER: 09/471,349
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: IN 3825/DEL/98
;  PRIOR FILING DATE: 1998-12-24
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;  LENGTH: 1245
;  TYPE: DNA
;  ORGANISM: Streptococcus equisimilis
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)...(1242)
US-09-940-235-1

Query Match             98.3%;
Best Local Similarity   99.0%;
Matches 1232; Mismatches 13;
```

Score 1224.2; DB 10; Length 1245;
Pred. No. 0;
Conservative   0; Indels   0; Gaps   0;

```
Qy   1081  ACCAACCGTATCATCAACCGTTTATATGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  ACCAACCGTATCATCAACCGTTTATATGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1140

Qy   1141  CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGACGAGAAGTTACAGCTACCTG 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGACGAGAAGTTACAGCTACCTG 1200

Qy   1201  CGTTATACAGGGACACCTATACCTGATAAACCTAACGACAAATAA 1245
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1201  CGTTATACAGGGACACCTATACCTGATAAACCTAACGACAAATAA 1245
```

RESULT 2
US-09-940-235-5
; Sequence 5, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-5

  Query Match   98.2%;   Score 1222.6;   DB 10;   Length 1377;
  Best Local Similarity 98.9%;   Pred. No. 0;
  Matches 1231;   Conservative 0;   Mismatches 14;   Indels 0;   Gaps 0;

```
Qy     1  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   133  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTGGTTGTT 192

Qy    61  AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   193  AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC 252

Qy   121  GACCTAACATCACGACCTGCTCATGGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253  GATCTAACATCACGACCTGCTCATGGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 312

Qy   181  AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTGAAAAAGCTGACTTACTA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   313  AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTGAAAAAGCTGACTTACTA 372

Qy   241  AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   373  AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 432

Qy   301  ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGTCTACTTTGCTGAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   433  ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGTCTACTTTGCTGAC 492

Qy   361  AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420
```

```
;   TITLE OF INVENTION:    PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
;   TITLE OF INVENTION:    CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
;   TITLE OF INVENTION:    PROTEIN
;   FILE REFERENCE: 07064-009002
;   CURRENT APPLICATION NUMBER: US/09/940,235
;   CURRENT FILING DATE:   2002-04-09
;   PRIOR APPLICATION NUMBER: 09/471,349
;   PRIOR FILING DATE: 1999-12-23
;   PRIOR APPLICATION NUMBER: IN 3825/DEL/98
;   PRIOR FILING DATE: 1998-12-24
;   NUMBER OF SEQ ID NOS: 28
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 1782
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hybrid cassette
US-09-940-235-11

  Query Match             98.2%;   Score 1222.6;   DB 10;   Length 1782;
  Best Local Similarity   98.9%;   Pred. No. 0;
  Matches 1231;   Conservative   0;   Mismatches   14;   Indels   0;   Gaps   0;

Qy       1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTGGTTGTT  597

Qy      61 AGCGGTGCTGCTCGTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     598 AGCGGTGCTGCTCGTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC  657

Qy     121 GACCTAACATCACGACCCTGCTCATGGAGGAAAGAGAGAGCTTAAGCCAAAATCA  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     658 GATCTAACATCACGACCCTGCTCATGGAGGAAAGACAGAGCTTAAGCCAAAATCA  717

Qy     181 AAACCATTTGCTACTGATAGTGGCGCGATGGCCACATAAACTTGAAAAAGCTGACTTACTA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     718 AAACCATTTGCTACTGATAGTGGCGCGATGTCAATAAACTTGAGAAAGCTGACTTACTA  777

Qy     241 AAGGCTATTCAAGAACAATTGATCGCTAAGTCGCAAGTCTCACTTTGAGGTC  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     778 AAGGCTATTCAAGAACAATTGATCGCTCCACGTCCCAGTAACGTAACGACTACTTTGAGGTC  837

Qy     301 ATTGATTTTGCAAGCGATGCAACCATTACTTACGAGGATAATCACGATGAC  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     838 ATTGATTTTGCAAGCGATGCAACCATTACTTACGAGGATAATCACGATGAC  897

Qy     361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     898 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  957

Qy     421 GTGCGCGTTAGACCATATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     958 GTGCGCGTTAGACCATATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  1017

Qy     481 GAATACTGTACAGTTTACTCCCTAACCCCTGATGACGATTTCAGACCAGGTCTCAAA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1018 GAATACTGTACAGTTTACTCCCTAACCCCTGATGACGATTTCAGACCAGGTCTCAAA  1077

Qy     541 GATACTAAGCTATTGAAAACACTAGCTATCGGTAGCCATCACCATCTCAAGAATTACTA  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1078 GATACTAAGCTATTGAAAACACTAGCTATCGGTAGCCATCACCATCTCAAGAATTACTA  1137

Qy     601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTATGAACGTGAC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1138 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTATGAACGTGAC  1197

Qy     661 TCCTCAATCGTCACTCACTGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1198 TCCTCAATCGTCACTCACTGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG  1257

Qy     721 TTTACTTACCATGTCAAAAATCGGAACATCTGGTCTG  780
```

```
Db     493 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  552

Qy     421 GTGCGCGTTAGACCATATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     553 GTGCGCGTTAGACCATATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  612

Qy     481 GAATACTGTACAGTTTACTCCCTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     613 GAATACTGTACAGTTTACTCCCTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  672

Qy     541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACGACCATCACCATCTCAAGAATTACTA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     673 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACGACCATCACCATCTCAAGAATTACTA  732

Qy     601 GCTCAAGCACAAAGCATTTAAACAAAAACCACCCAGGCTATACGCTATACGTGAC  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     733 GCTCAAGCACAAAGCATTTAAACAAAAACCACCCAGGCTATACGCTATACGTGAC  792

Qy     661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     793 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG  852

Qy     721 TTTACTTACCATGTCAAAAATCGGAACAGCTTATGAGATCAATAAAAAATCTGGTCTG  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     853 TTTACTTACCGTGTTAAAAATCGGGAACAAGTTATAGGATCAATAAAAAATCTGGTCTG  912

Qy     781 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     913 AATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG  972

Qy     841 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGT  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     973 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  1032

Qy     901 GATGTCAACACCAACGAAATTGCTAAAAAGCGAGCAGCTCCTTAACGACTAGCGGAACGTAAC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1033 GATGTCGATACCAACGAAATTGCTAAAAAGTGAGCAGCTCCTTAACGACTAGCGGAACGTAAC  1092

Qy     961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTCACAACAATCTC  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1093 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTCACAACAATCTC  1152

Qy    1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC  1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1153 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC  1212

Qy    1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGGCGACCCGAAGGAGAGAATGCTAGCTAT  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1213 ACCAACCGTATCATAACCGTTTATATGGGCAAGGCGACCCGAAGGAGAGAATGCTAGCTAT  1272

Qy    1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1273 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  1332

Qy    1201 CGTTATACAGGGACGACCTATACCTGATAACCCTAACGACAAATAA  1245
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    1333 CGTTATACAGGGACGACCTATACCTGATAACCCTAACGACAAATAA  1377
```

```
RESULT 3
US-09-940-235-11
;   Sequence 11, Application US/09940235
;   Publication No. US20030059921A1
;   GENERAL INFORMATION:
;   APPLICANT: Kumar, Rajesh
;   APPLICANT: Sahni, Girish
;   APPLICANT: Roy, Chait
;   APPLICANT: Rajagopal, Kammara
;   APPLICANT: Nihalani, Deepak
;   APPLICANT: Sundaram, Vasudha
;   APPLICANT: Yadav, Mahavir
;   TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
```

```
Db    83  ATAGCTGGTCCTGAATGGCTACTAGATGTCCTTCTGTAAATAACAGCCAATTCGTTGTT  142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    61  AGCGTTGCTGGTACTGTTGAGGGGACGAATCAGACATTAGTCTTAAATTTTTGAAATT  120
           ||| ||||||||||||||||||||||||||| |||||||||||||| |||||||||
Db   143  AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC  202

Qy   121  GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180
           || |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203  GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  262

Qy   181  AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAAACTTGAAAAAGCTGACTTACTA  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263  AAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAAACTTGAAAAAGCTGACTTACTA  322

Qy   241  AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323  AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  382

Qy   301  ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   383  ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC  442

Qy   361  AAAGATGGTTCGGTAACCTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   443  AAAGATGGTTCGGTAACCTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT  502

Qy   421  GTGCGCGTTAGACCATATAAAGAAAAACCAAATACAAAATCAAGCGAAATCTGTTGATGTG  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   503  GTGCGCGTTAGACCATATAAAGAAAAACCAAATACAAAACAAGCGAAATCTGTTGATGTG  562

Qy   481  GAATATACTGTACAGTTTACTCCCTAAACCTGATGACGATTTCAGACACCAGGTCTCAAA  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   563  GAATATACTGTACAGTTTACTCCCTAAACCTGATGACGATTTCAGACCAGGTCTCAAA  622

Qy   541  GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   623  GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA  682

Qy   601  GCTCAAGCACAAAGCATTTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   683  GCTCAAGCACAAAGCATTTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC  742

Qy   661  TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   743  TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG  802

Qy   721  TTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   803  TTTTACTTACCGTGTTAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  862

Qy   781  AATGAAGAAATAAACAACACTGACCTTATCTCTGAGAAAATATTACGTCCTTAAAAAGGG  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   863  AATGAAGAAATAAACAACACTGACCTTATCTCTGAGAAAATATTACGTCCTTAAAAAGGG  922

Qy   841  GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   923  GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  982

Qy   901  GATGTCAACACCAACGAATTGCTAAAAGTGCTAAAAGTGGCAGCTCTTAACAGCTAGCGAACGTAAC  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   983  GATGTCGATACCAACGAATTGCTAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC  1042

Qy   961  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC  1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1043  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC  1102

Qy  1021  GATGCTTTTGGTATTATGGGACTATACCTTAACTGAGAAAAGTAGAGGATAATCACGATGAC  1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1103  GATGCTTTTGGTATTATGGGACTATACCTTAACTGAGAAAAGTAGAGGATAATCACGATGAC  1162

Qy  1081  ACCAACCGTATCATAACCGTTCATCATGGGCAAGCAGCCCGGAAGGAGAGAATGCTAGCTAT  1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1163  ACCAACCGTATCATAACCGTTCATCATGGGCAAGCAGCCCGGAAGGAGAGAATGCTAGCTAT  1222

Db  1258  TTTACTTACCGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG  1317
           ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||

Qy   781  AATGAAGAAATAAACAACACTGACCTTATCTCTGAGAAAATATTACGTCCTTAAAAAGGG  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1318  AATGAAGAAATAAACAACACTGACCTTATCTCTGAGAAAATATTACGTCCTTAAAAAGGGG  1377

Qy   841  GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1378  GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  1437

Qy   901  GATGTCAACACCAACGAATTGCTAAAAGTGCTAAAAGTGGCAGCTCTTAACAGCTAGCGAACGTAAC  960
           |||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1438  GATGTCGATACCAACGAATTGCTAAAAGTGCTAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC  1497

Qy   961  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC  1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1498  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC  1557

Qy  1021  GATGCTTTTGGTATTATGGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC  1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1558  GATGCTTTTGGTATTATGGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC  1617

Qy  1081  ACCAACCGTATCATAACCGTTCATCATGGGCAAGCGACCCGGAAGGAGAATGCTAGCTAT  1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1618  ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGGAAGGAGAATGCTAGCTAT  1677

Qy  1141  CATTAGCCTATGATAAAGATCGTTATACCGGAAGAACGAGAAGTTTACAGCTACCTG  1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1678  CATTTAGCCTATGATAAAGATCGTTATACCGGAAGAACGAGAAGTTTACAGCTTACCTG  1737

Qy  1201  CGTTATACAGGACACCTATACCTGATAACCTAACGACAAATAA  1245
           ||||||||||||||||||||||||||||||||||||||||||
Db  1738  CGTTATACAGGACACCTATACCTGATAACCCTAACGACAAATAA  1782
```

RESULT 4
US-09-940-235-6

; Sequence 6, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERRED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-6

Query Match   97.2%;   Score 1209.8;   DB 10;   Length 1327;
Best Local Similarity 98.2%;   Pred. No. 0;
Matches 1223;   Conservative   0;   Mismatches   22;   Indels   0;   Gaps   0;

```
Qy    1  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTAGTTGTT   60
          || ||||| |||||||||||||||||||| ||||||||| ||||||||||||||||||
```

```
Qy  1141 CATTTAGCCTATGATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGAGTTTACAGCTACCTG 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1223 CATTTAGCCTATGATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGAGTTTACAGCTACCTG 1282

Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1283 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1327
```

RESULT 5
US-09-940-235-10
; Sequence 10, Application US/09940235
; Publication No. US2003005992lA1
; GENERAL INFORMATION:
;   APPLICANT: Kumar, Rajesh
;   APPLICANT: Sahni, Girish
;   APPLICANT: Roy, Chait
;   APPLICANT: Rajagopal, Kammara
;   APPLICANT: Nihalani, Deepak
;   APPLICANT: Sundaram, Vasudha
;   APPLICANT: Yadav, Mahavir
;   TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
;   TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
;   TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
;   TITLE OF INVENTION: PROTEIN
;   FILE REFERENCE: 07064-009002
;   CURRENT APPLICATION NUMBER: US/09/940,235
;   CURRENT FILING DATE: 2002-04-09
;   PRIOR APPLICATION NUMBER: 09/471,349
;   PRIOR FILING DATE: 1999-12-23
;   PRIOR APPLICATION NUMBER: IN 3825/DEL/98
;   PRIOR FILING DATE: 1999-12-24
;   NUMBER OF SEQ ID NOS: 28
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 10
;   LENGTH: 1661
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hybrid cassette
US-09-940-235-10

Query Match   90.6%;   Score 1127.8;   DB 10;   Length 1661;
Best Local Similarity   98.5%;   Pred. No. 1.5e-303;
Matches 1138;   Conservative   0;   Mismatches 17;   Indels 0;   Gaps 0;

Qy   1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   183 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTGGTTGTT 242

Qy   61 AGCGGTTGCTGTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   243 AGCGGTTGCTGTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATC 302

Qy   121 GACCTAACATCACGACCTGCTCATGGAGGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
         || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 GATCTAACATCACGACCTGCTCATGGAGGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 362

Qy   181 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA 240
         |||||||||||||||||||||||||||||| ||||||||||| ||||||||||||||||
Db   363 AAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAACTTGAGAAAGCTGACTTACTA 422

Qy   241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   423 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 482

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCAGAAACGGCAAGGTCTTACTTTGCTGAC 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   483 ATTGATTTTGCAAGCGATGCAACCATTACTGATCAGAAACGGCAAGGTCTTACTTTGCTGAC 542

Qy   361 AAAGATGGTTCGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   543 AAAGATGGTTCGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 602

Qy   421 GTGCGCGGTTAGACCATATAAAGAAGAAAAACCAATACAAATCAAGCGAAATCGTTGATGTG 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   603 GTGCGCGGTTAGACCATATAAAGAAGAAAAACCAATACAAATCAAGCGAAATCGTTGATGTG 662

Qy   481 GAATATACTGTACAGTTACTCCCTTAAACCTGATGACGATTCAGACGGTCTCAAA 540
         |||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   663 GAATATACTGTACAGTTACTCCCTTAAACCTGATGACGATTCAGACCAGGTCTCAAA 722

Qy   541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   723 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 782

Qy   601 GCTCAAGCACAAAGCATTTAAACAAAACCCACCCAGGCTATACGTATACGAACGTGAC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   783 GCTCAAGCACAAAGCATTTAAACAAAACCCACCCAGGCTATACGTATACGAACGTGAC 842

Qy   661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   843 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG 902

Qy   721 TTTACTTACCATGTCAAAAATCGGGACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780
         |||||||||| ||||| |||||||||||||||||| ||||||||||||||||||||||
Db   903 TTTACTTACCGTGTAAAAATCGGGACAAGCTTATAGGATCAATAAAAATCTGGTCTG 962

Qy   781 AATGAAGAAATAAACAACACTGACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   963 AATGAAGAAATAAACAACACTGACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1022

Qy   841 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTGCAGTCGTTCACCATCAAATACGTT 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1023 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTGCAGTCGTTCACCATCAAATACGTT 1082

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGTCTTAACAGCTAGCGAACGTAAC 960
         ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1083 GATGTCGATACCAACGAATTGCTAAAAAGCGAGCAGTCTTAACAGCTAGCGAACGTAAC 1142

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGCTGAAACTACTCTACAACAATCTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1143 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGCTGAAACTACTCTACAACAATCTC 1202

Qy 1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1203 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1262

Qy 1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1263 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1322

Qy 1141 CATTTAGCCTATGAT 1155
         ||||||||| |||||
Db 1323 CATTTAGCCGGTGGT 1337
```

RESULT 6
US-09-940-235-12
; Sequence 12, Application US/09940235
; Publication No. US2003005992lA1
; GENERAL INFORMATION:
;   APPLICANT: Kumar, Rajesh
;   APPLICANT: Sahni, Girish
;   APPLICANT: Roy, Chait
;   APPLICANT: Rajagopal, Kammara
;   APPLICANT: Nihalani, Deepak
;   APPLICANT: Sundaram, Vasudha
;   APPLICANT: Yadav, Mahavir
;   TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
;   TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
;   TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
;   TITLE OF INVENTION: PROTEIN
;   FILE REFERENCE: 07064-009002
;   CURRENT APPLICATION NUMBER: US/09/940,235

```
;  CURRENT FILING DATE: 2002-04-09
;  PRIOR APPLICATION NUMBER: 09/471,349
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: IN 3825/DEL/98
;  PRIOR FILING DATE: 1998-12-24
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 12
;  LENGTH: 2096
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Hybrid cassette
US-09-940-235-12

Query Match             90.3%;  Score 1124.6;  DB 10;  Length 2096;
Best Local Similarity   98.4%;  Pred. No. 1.3e-302;
Matches 1136;  Conservative  0;  Mismatches 19;  Indels  0;  Gaps  0;

Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   588 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTGGTTGTT 647

Qy    61 AGCGTTGCTGTGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTTGAAATT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   648 AGCGTTGCTGTGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTTGAAATC 707

Qy   121 GACCTAACATCACGACCTGCTCATGGAAGACAGAGCAAGGCTAAGTCCAAAATCA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   708 GATCTAACATCACGACCTGCTCATGGAAGACAGAGACAGAGCAAGGCTAAGTCCAAAATCA 767

Qy   181 AAACCATTTGCTACTGATAGTGGCGCGATGCGCGCGATGTCACATAAACTGAAAAAGTGACTTACTA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   768 AAACCATTTGCTACTGATAGTGGCGCGATGTGGCGCGATGTCACATAAACTGAGAAAGTGACTTACTA 827

Qy   241 AAGGCTATTCAAGAGAACAAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   828 AAGGCTATTCAAGAGAACAAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 887

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTCTACTTTGCTGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   888 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTTCTACTTTGCTGAC 947

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   948 AAAGATGGTTCGGTAACCTTGCCGACCCAACCGTCCAAGAATTTTTGCTAAGCGGACAT 1007

Qy   421 GTGCGCGTTAGACCATATAAAGAAAACCAATACAAAATCAAGCGAAATCGTTGATGTG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1008 GTGCGCGTTAGACCATATAAAGAAAACCAATACAAAATCAAGCGAAATCGTTGTGATGTG 1067

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCCAGGTCTCAAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1068 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCCAGGTCTCAAA 1127

Qy   541 GATACTAAGCTATTGAAAACACTAGCTAGCTATCGGTGACCACCATCAAGAATTACTA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1128 GATACTAAGCTATTGAAAACACTAGCTAGCTATCGGTGACCACCATCAAGAATTACTA 1187

Qy   601 GCTCAAGCACAAAAGCATTTAAACAAAACCCAGGCTATACGATTTATGAACGTAGAC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1188 GCTCAAGCACACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAGAC 1247

Qy   661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1248 TCCTCAATCGTCACTCGTCACTCATGACAATGACATTTTCCGTACGATTTTTACCAATGGATCAAGAG 1307

Qy   721 TTTACTTACCATGTCAAAAATCTGGGAACAAGCTTATGAGATCAATAAAAAAATCTCGGTCTG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1308 TTTACTTACCGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAAATCTCGGTCTG 1367

Qy   781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCGTCCTTAAAAAAGGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1368 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCGTCCTTAAAAAAGGG 1427

Qy   841 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTCGAGAAACTGTTCACCATCAAATAGTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1428 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTCGAGAAACTGTTCACCATCAAATACGTT 1487

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1488 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1547

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1548 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1607

Qy  1021 GATGCTTTGGTATTATGGACTATACCTTAACTGAAAAAGTAGAGGATAATCACGATGAC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1608 GATGCTTTGGTATTATGGACTATACCTTAACTGAAAAAGTAGAGGATAATCACGATGAC 1667

Qy  1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1668 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAC 1727

Qy  1141 CATTTAGCCTATGAT 1155
           ||||||||||||||||
Db  1728 CATTTAGCTGGTGGT 1742
```

RESULT 7
US-09-940-235-9
```
;  Sequence 9, Application US/09940235
;  Publication No. US20030059921A1
;  GENERAL INFORMATION:
;     APPLICANT:  Kumar, Rajesh
;     APPLICANT:  Sahni, Girish
;     APPLICANT:  Roy, Chait
;     APPLICANT:  Rajagopal, Kammara
;     APPLICANT:  Nihalani, Deepak
;     APPLICANT:  Sundaram, Vasudha
;     APPLICANT:  Yadav, Mahavir
;     TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
;     TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
;     TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
;     TITLE OF INVENTION: PROTEIN
;  FILE REFERENCE: 07064-009002
;  CURRENT APPLICATION NUMBER: US/09/940,235
;  CURRENT FILING DATE: 2002-04-09
;  PRIOR APPLICATION NUMBER: 09/471,349
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: IN 3825/DEL/98
;  PRIOR FILING DATE: 1998-12-24
;  NUMBER OF SEQ ID NOS: 28
;  SEQ ID NO 9
;  LENGTH: 1541
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Hybrid cassette
US-09-940-235-9

Query Match             89.3%;  Score 1111.8;  DB 10;  Length 1541;
Best Local Similarity   97.7%;  Pred. No. 4.2e-299;
Matches 1128;  Conservative  0;  Mismatches 27;  Indels  0;  Gaps  0;

Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    33 ATAGCTGGTGTCCTGAATGGCTACTAGACGTGCTCCTTCTGTAAATAACAGCCAATTGGTTGTT 92

Qy    61 AGCGGTTGCTGTGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTTGAAATT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    93 AGCGGTTGCTGTGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTTGAAATC 152

Qy   121 GACCTAACATCACGACCTGCTCATGGAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

US-10-369-493-29120/c
; Sequence 29120, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 29120
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-10-369-493-29120

Query Match   3.5%; Score 43.4; DB 16; Length 556;
Best Local Similarity   60.7%; Pred. No. 0.21;
Matches   71; Conservative   0; Mismatches   46; Indels   0; Gaps   0;

Qy   688 ATTTTCCGTACGATTTACCAAATGATCAAGAGTTACTTACCATGTCAAAAATCGGGAA 747
          ||||||||||||||||||| ||||||| ||||||||||||| ||||||||||||||
Db   292 ATTTTCAGAATGTTTTTCGAATTCATCAAACATTTTTACCGATCATCAAAAAACAGCAA 233

Qy   748 CAAGCTTATGAGATCAATAAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGAC 804
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   232 AAAAATTTAAGTTACCAAACTATTTCCCTCAGAAAAAAACTTACAACAACAATCAC 176

RESULT 9
US-10-221-613-270/c
; Sequence 270, Application US/10221613
; Publication No. US2004002912A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
;   DE 10013847.00
;   DE 10019058.8
;   DE 10019173.8
;   DE 10032529.7
;   DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
;   2000-03-15
;   2000-04-06
;   2000-04-07
;   2000-06-30
;   2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 270
; LENGTH: 5504
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-270

Query Match   3.4%; Score 42; DB 13; Length 5504;
Best Local Similarity   46.8%; Pred. No. 2;
Matches 132; Conservative   0; Mismatches 132; Indels   0; Gaps   0;

```
Db         ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       153 GATCTAACATCACGACCTGCTCATGGAGCAAGAAACAGCAAGCTTAAGTCAAAATCA 212

Qy     181 AAACCATTTGCTACTGACTAGTGGCGGATGCCCATAAAACTTGAAAAAGCTGACTTACTA 240
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db     213 AAACCATTTGCTACTGATAGTGGCGGATGTCACATAAAACTTGAGAAAGCTGACTTACTA 272

Qy     241 AAGGCTATTCAAGAACAATTGATCGCTACGTCCACAGTAACGACGACTACTTTGAGGTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 AAGGCTATTCAAGAACAATTGATCGCTACGTCCACAGTAACGACGACTACTTTGAGGTC 332

Qy     301 ATTGATTTTGCAAGCGATGCAACCATTACTGACTTGATCGAAAACGGCAAGGTCTCACTTTGCTGAC 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     333 ATTGATTTTGCAAGCGATGCAACCATTACTGACTTGATCGAAACGGCAAGGTCTCACTTTGCTGAC 392

Qy     361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     393 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT 452

Qy     421 GTGCGCGGTTAGACCATATAAAGAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     453 GTGCGCGGTTAGACCATATAAAGAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 512

Qy     481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCCAGGTCTCAAA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     513 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCCAGGTCTCAAA 572

Qy     541 GATACTAAGCTATTGAAAAACACTAGCTATTCGGTGCACACCATCACATCTCAAGAATTACTA 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     573 GATACTAAGCTATTGAAAAACACTAGCTATTCGGTGCACACCATCACATTTTTACCGATCATCAAGAATTACTA 632

Qy     601 GCTCAAGCACAAAGCACATTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     633 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC 692

Qy     661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     693 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752

Qy     721 TTTACTTACCACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     753 TTTACTTACCGGTGTTAAAAATCGGGAACAAGCTTATAGAGTCAATAAAAAATCTGGTCTG 812

Qy     781 AATGAAGAAATAAACAACAACACTGACCTCTGATCTCTGAGAAAATATTACGTCCTTAAAAAGGG 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     813 AATGAAGAAATAAACAACAACACTGACCTCTGATCTCTGAGAAAATATTACGTCCTTAAAAAGGG 872

Qy     841 GAAAAGCCGTATGATCCCCTTTGATCGCAGTCAGTTCACCTGAAACTGTTCACCATCAAATACGTT 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     873 GAAAAGCCGTATGATCCCCTTTGATCGCAGTCAGTTCACCTGAAACTGTTCACCATCAAATACGTT 932

Qy     901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     933 GATGTCGATACCAACGAATTGCTAAAAAGTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 992

Qy     961 TTAGACTTCAGAGATTTATACGATCCTCGTGATGAAGGCTAAACTACTCTCCACAACAATCTC 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     993 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAAGGCTAAACTACTCTCCACAACAATCTC 1052

Qy    1021 GATGCTTTTTGGTATTATGGACTATACCTTAACTGAAAGTAGAGGGATAATCACGATGAC 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1053 GATGCTTTTTGGTATTATGGACTATACCTTAACTGAAAGTAGAGGGATAATCACGATGAC 1112

Qy    1081 ACCAACCGTATCATAACCGTTTATATGGCAAGCGACCCGAAGAGGAGAATGCTAGCTAT 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1113 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAC 1172

Qy    1141 CATTTAGCCTATGAT 1155
           |||||||||||||||
Db    1173 CATTTAGCTGCTGGT 1187
```

RESULT 8

Qy   538 AAAGATACTAGCTAGTGAATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTA 597
          |||||| |||||||||||
Db 1317 AAAAATATAATCAAACTAACAATCCTTTTAAAATCTTTTCTCTTACTAATATCAAGAAAAATA 1258

Qy   598 CTAGCTCAAGCACAAAGCATTTAAACAAAACCACCAGGCTATACGATTTATGAACGT 657
          ||||||| ||||||
Db 1257 CCACCGAAACAAATAAATACTTTAAAATAAAAAACAAAAACTCTCTAAATAAAAAACTT 1198

Qy   658 GACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAA 717
          |||||| ||||||
Db 1197 CAAAAAAAAATACAAACGACCTAAAAAAATATACTACGATCAAATTTAATCCTAAAAAAA 1138

Qy   718 GAGTTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGT 777
          ||||||| ||||||
Db 1137 AATAATCCTAAAAAACAAATAAAAAAATAAAAAAAAAATATAAAAAAAAAAAACGCT 1078

Qy   778 CTGAATGAAGAAATAAACAACACTGCTGACCTCTCTGAGAAA 819
          |||||| |||||
Db 1077 AATATAAAAAAAAAAAAAAAAAACGAAAATAAAAAACCTAAAAA 1036

RESULT 10
US-10-311-455-340/c
; Sequence 340, Application US/10311455
; Publication No. US2003014366A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 340
; LENGTH: 6071
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-340

Query Match   3.4%; Score 42; DB 15; Length 6071;
Best Local Similarity 46.3%; Pred. No. 2.1;
Matches 138; Conservative   0; Mismatches 160; Indels   0; Gaps   0;

Qy   551 TATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTAGCTCAAGCAC 610
          |||||| |||||||
Db 2019 TATAAAAAACTAATTCTAAAATAAAATCATTTATTCTAAAATATAAAATATAAAAAAA 1960

Qy   611 AAAGCATTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGACTCCTCAATCG 670
          ||||| |||| |||
Db 1959 AAACCACCAATCCAAACTAATAAAAACGAAAAAATAATAAATACCTACCTCAACTT 1900

Qy   671 TCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACC 730
          || | |||||
Db 1899 AAACTAAATAAATATAAAAACTAAACAATAAAAAATATATAAATCTGGTCGAATGAAGAAA 1840

Qy   731 ATGTCAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCGAATGAAGAAA 790
          ||| ||
Db 1839 TAAAAAAAATCCCGATTAAAATAAAAACAAAAAAAAAATCTAAAAAAAAAA 1780

Qy   791 TAAACAACACTGACCTGATCTCTGAGAAATAAATATTACGTCCTTAAAAAGGGGAAAAGCC 848
          |||||| |||
Db 1779 AACCAAAAACTAAAAATCCCAACTCCGAAAATAAAAAATATAAAAAACGAAAAAAACC 1722

RESULT 11
US-10-240-485-36/c
; Sequence 36, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: Metastasis
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 36
; LENGTH: 6071
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-36

Query Match   3.4%; Score 42; DB 15; Length 6071;
Best Local Similarity 46.3%; Pred. No. 2.1;
Matches 138; Conservative   0; Mismatches 160; Indels   0; Gaps   0;

Qy   551 TATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTAGCTCAAGCAC 610
          |||||| |||||||
Db 2019 TATAAAAAACTAATTCTAAAATAAAATCATTTATTCTAAAATATAAAATATAAAAAAA 1960

Qy   611 AAAGCATTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGACTCCTCAATCG 670
          ||||| |||| |||
Db 1959 AAACCACCAATCCAAACTAATAAAAACGAAAAAATAATAAATACCTACCTCAACTT 1900

Qy   671 TCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACC 730
          || | |||||
Db 1899 AAACTAAATAAATATAAAAACTAAACAATAAAAAATATATAAATCTGGTCGAATGAAGAAA 1840

Qy   731 ATGTCAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCGAATGAAGAAA 790
          ||| ||
Db 1839 TAAAAAAAATCCCGATTAAAATAAAAACAAAAAAAAAATCTAAAAAAAAAA 1780

Qy   791 TAAACAACACTGACCTGATCTCTGAGAAATAAATATTACGTCCTTAAAAAGGGGAAAAGCC 848
          |||||| |||
Db 1779 AACCAAAAACTAAAAATCCCAACTCCGAAAATAAAAAATATAAAAAACGAAAAAAACC 1722

RESULT 12
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence

```
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (3294164)
US-10-312-841-1

Query Match             3.4%;  Score 42;  DB 15;  Length 3673778;
Best Local Similarity   51.2%;  Pred. No. 91;
  Matches  124;  Conservative    0;  Mismatches  115;  Indels    3;  Gaps    1;

Qy        763 AATAAAAATCTGGTCTCGAATGAAGAGAAATAAACAACACTGACCTGATCTCTGAGAAATAT 822
              |||   ||||||||  ||| |||| |||||| ||| ||||||||| |||| |||||||||||
Db    1605592 AATATCAAATCTACTTATATTTAAAAAATAATAAAACTTACCTCTTAATAAAAAAAAATAA 1605633

Qy        823 TACGTCCTTAAAAAAGGGGAAAAGCCGTATGATCCCTTGATCGCAGTCACTGAAACTG 882
              ||| ||||||| ||   ||||||||||   ||||     ||||||| |||||  |||||
Db    1605632 TAAATC---AAACTAAAAAAAAAACAAATACGATAAAAAATATTTAAATATTTTTACAACC 1605576

Qy        883 TTCACCATCAAATACGTTGATGTCAACACCAACGAATTGCTAAAAAGGCGAGCAGCTCTTA 942
              ||||||||||||||  |||||||||| || |||| |||| ||||||||  |||||||||||
Db    1605575 CTCTTTAAAAAATACATCAAATTCAAAACATCCTATCTACCAAAAAAAAAAACACCATT 1605516

Qy        943 ACAGCTAGCGAACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAA 1002
              ||| | |  ||||    |||||| |||||||||||||||||||||| ||||| ||| |
Db    1605515 ACGCCTTACCCCTCCAATTTAAACCCTATAATTCTATAATATCCTATAAAAAAATCATAT 1605456

Qy       1003 CT 1004
              ||
Db    1605455 CT 1605454


RESULT 13
US-10-311-455-658/c
;   Sequence 658, Application US/10311455
;   Publication No. US20030143606A1
;   GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
;   TITLE OF INVENTION: cytosine methylation
;   FILE REFERENCE: 5013.1014
;   CURRENT APPLICATION NUMBER: US/10/311,455
;   CURRENT FILING DATE: 2002-12-16
;   PRIOR APPLICATION NUMBER: PCT/EP01/07537
;   PRIOR FILING DATE: 2001-07-02
;   PRIOR APPLICATION NUMBER: DE 10032529.7
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: DE 10043826.1
;   PRIOR FILING DATE: 2000-09-01
;   NUMBER OF SEQ ID NOS: 2424
;   SEQ ID NO 658
;   LENGTH: 6392
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-658

Query Match             3.4%;  Score 41.8;  DB 15;  Length 6392;
Best Local Similarity   50.2%;  Pred. No. 2.5;
  Matches  103;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy        546 TAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTAGCTCA 605
              ||||||||||||||||| |||| ||||||||||||||||||||| || ||||||||||||
Db        377 TAACTTACTAAAAAACCCTTCCTAACGAACTAAACCTAACTCTTCCTAAACTAAATCCTAA 318

Qy        606 AGGACAAAGCATTTTAAACAAAACCCACCCCAGGCTATACGATTTATGAACGTGACTCCTC 665
              |  |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||
Db        317 ACTAACACACTAATAACTCTAAACTCTTCCCGATCAAATTATAACAAAACTTTTACCCATA 258

Qy        666 AATCGTCACTCATGACAATGACAATGCATTTTCCGTACGATTTTCCAATGGATCAAGAGTTTAC 725
```

```
Db        257 AATAAATTTCAAACGAAAATATCAAAAATCCTTACCGATATCCTAAAAATTACAAAATAAATC 198
              |||   |||   ||   |||   |||   ||  |   |||

Qy        726 TTACCATGTCAAAAATCGGGAACAA 750
              ||| ||||  |||
Db        197 TTATCATAAAAAAATTCTAAATCTA 173


RESULT 14
US-10-240-485-60/c
;   Sequence 60, Application US/10240485
;   Publication No. US20030148327A1
;   GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with
;   TITLE OF INVENTION: Metastasis
;   FILE REFERENCE: 5013.1007
;   CURRENT APPLICATION NUMBER: US/10/240,485
;   CURRENT FILING DATE: 2002-10-02
;   PRIOR APPLICATION NUMBER: PCT/EP01/03970
;   PRIOR FILING DATE: 2001-04-06
;   PRIOR APPLICATION NUMBER: DE 10019058.8
;   PRIOR FILING DATE: 2000-04-06
;   PRIOR APPLICATION NUMBER: DE 10019173.8
;   PRIOR FILING DATE: 2000-04-07
;   PRIOR APPLICATION NUMBER: DE 10032529.7
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: DE 10043826.1
;   PRIOR FILING DATE: 2000-09-01
;   NUMBER OF SEQ ID NOS: 202
;   SEQ ID NO 60
;   LENGTH: 6392
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-60

Query Match             3.4%;  Score 41.8;  DB 15;  Length 6392;
Best Local Similarity   50.2%;  Pred. No. 2.5;
  Matches  103;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy        546 TAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTAGCTCA 605
              ||||||||||||||||| |||| ||||||||||||||||||||| || ||||||||||||
Db        377 TAACTTACTAAAAAACCCTTCCTAACGAACTAAACCTAACTCTTCCTAAACTAAATCCTAA 318

Qy        606 AGGACAAAGCATTTTAAACAAAACCCACCCCAGGCTATACGATTTATGAACGTGACTCCTC 665
              |  |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||
Db        317 ACTAACACACTAATAACTCTAAACTCTTCCCGATCAAATTATAACAAAACTTTTACCCATA 258

Qy        666 AATCGTCACTCATGACAATGACAATATCAAAAATCCTTACCGATATCCTAAAAATTACAAAATAAATC 198
              |||   |||   ||   |||   |||   ||  |   |||
Db        257 AATAAATTTCAAACGAAAATATCAAAAATCCTTACCGATATCCTAAAAATTACAAAATAAATC 198

Qy        726 TTACCATGTCAAAAATCGGGAACAA 750
              ||| ||||  |||
Db        197 TTATCATAAAAAAATTCTAAATCTA 173


RESULT 15
US-10-311-455-61/c
;   Sequence 61, Application US/10311455
;   Publication No. US20030143606A1
;   GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determi
;   TITLE OF INVENTION: cytosine methylation
;   FILE REFERENCE: 5013.1014
;   CURRENT APPLICATION NUMBER: US/10/311,455
```

```
;   CURRENT FILING DATE: 2002-12-16
;   PRIOR APPLICATION NUMBER: PCT/EP01/07537
;   PRIOR FILING DATE: 2001-07-02
;   PRIOR APPLICATION NUMBER: DE 10032529.7
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: DE 10043826.1
;   PRIOR FILING DATE: 2000-09-01
;   NUMBER OF SEQ ID NOS: 2424
;   SEQ ID NO 61
;      LENGTH: 6283
;      TYPE: DNA
;      ORGANISM: Artificial Sequence
;      FEATURE:
;      OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-61

Query Match       3.3%;   Score 41.6;  DB 15;  Length 6283;
Best Local Similarity  46.3%;   Pred. No. 2.8;
Matches  137;   Conservative    0;  Mismatches  159;  Indels    0;  Gaps     0;

Qy      677 ATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCATGTCA 736
                |||||||||  |||      |||   ||  |||||    |||  |||||||||
Db      391 ATAACTTACACATATTTTAAATTTCATTAAATTTAAAATAACAAATTTACTTAAATATAAT 332

Qy      737 AAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACA 796
                |  |||| ||   ||      |||||| ||   ||||  ||   |||| ||||
Db      331 CAAACATATATAAAACATAAATTTCTACAAAAAAACTTTCTCAAAATACAAACCTAAAAA 272

Qy      797 ACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGGGAAAAGCCGTATGATC 856
                ||||  ||||||| |||||     |||||| |||||  ||||||
Db      271 AAACATAATTCCTAAATCCTAAATATAATAACCCAAAAAAATTAACATTCCCTTCTCCAC 212

Qy      857 CCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTTGATGTCAACACCAACG 916
                  ||   ||||      |||||     ||||||  |||  ||||||
Db      211 TAAATTATACCTCTACGACAAAACTCTATCTTATCAATTCCACTATATTCCTAAATTAAAA 152

Qy      917 AATTGCTAAAAAGCGAGCAGCTCCTTAACAGCTAGCGAACGTAACTTAGACTTCAGA 972
                |||| |||      |||||   ||   || ||   |||||  || || |||
Db      151 AATTCTTTAAAACCAATAAAAACAAATAAAACACAACATAACAAAATCCTAAAACACA 96
```

Search completed: May 6, 2004, 08:06:01
Job time : 1271 secs

```
                  GenCore version 5.1.6
       Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    May  5, 2004, 23:29:56 ; Search time 552 Seconds
                                    (without alignments)
                              9581.534 Million cell updates/sec

Title:          US-09-882-509-3
Perfect score:  1245
Sequence:       1 attgctggacctgagtggct.........ataccctaacgacaaataa 1245

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:      6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  N Geneseq_29Jan04:*
            1: geneseqn1980s:*
            2: geneseqn1990s:*
            3: geneseqn2000s:*
            4: geneseqn2001as:*
            5: geneseqn2001bs:*
            6: geneseqn2002s:*
            7: geneseqn2003as:*
            8: geneseqn2003bs:*
            9: geneseqn2003cs:*
           10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1241.8	99.7	1245	2	AAQ20665	Aaq20665 SKC-2 str
2	1241.8	99.7	1335	2	AAQ12156	Aaq12156 Streptoki
3	1241.8	99.7	1458	2	AAQ12162	Aaq12162 Factor Xa
4	1241.8	99.7	1512	2	AAQ12158	Aaq12158 Streptoki
5	1241.8	99.7	2589	2	AAQ12160	Aaq12160 OmpAL str
6	1238.0	99.5	1467	2	AAQ12490	Aaq12490 Factor Xa
7	1224.2	98.3	1245	3	AAA37633	Aaa37633 S. equisi
8	1224.2	98.3	1473	2	AAQ05603	Aaq05603 Streptoki
9	1222.6	98.2	1377	3	AAA37622	Aaa37622 Streptoki
10	1222.6	98.2	1782	3	AAA37642	Aaa37642 Chimeric
11	1222.6	98.2	2030	2	AAQ11651	Aaq11651 FB-FB-SK
12	1221.2	98.1	1242	2	AAX80492	Aax80492 Streptoco
13	1221.2	98.1	1254	6	ABA05546	Aba05546 Streptoki
14	1221.2	98.1	2385	2	AAX80497	Aax80497 Streptoki
15	1221.2	98.1	8893	6	ABA05547	Aba05547 Maxadilan
16	1219.6	98.0	1242	2	AAX16632	Aax16632 Streptoco
17	1219.4	97.9	1323	2	AAT29961	Aat29961 Vector pS
18	1216.4	97.7	1242	2	AAX16633	Aax16633 Streptoco
19	1210.2	97.2	2566	2	AAT77778	Aat77778 Coding se
20	1209.8	97.2	1327	3	AAA37628	Aaa37628 Streptoki
21	1206.4	96.9	1209	3	AAZ99249	Aaz99249 DNA encod
22	1203.8	96.7	1245	3	AAZ99250	Aaz99250 DNA encod
23	1192.2	95.8	1407	1	AAN70106	Aan70106 DNA encod
24	1182.6	95.0	2568	1	AAN50493	Aan50493 Sequence
25	1170.4	94.0	1242	5	AAFB2144	Aafb2144 Mutant st
26	1127.8	90.6	1661	3	AAA37637	Aaa37637 Chimeric
27	1124.6	90.3	2096	3	AAA37643	Aaa37643 Chimeric
28	1116	89.6	1122	3	AAZ99251	Aaz99251 DNA encod
29	1116	89.6	1158	3	AAZ99252	Aaz99252 DNA encod
30	1111.8	89.3	1541	3	AAA37644	Aaa37644 Chimeric
31	1103.4	88.6	1119	2	AAQ12159	Aaq12159 Truncated
32	1103.2	88.6	2253	2	AAQ12161	Aaq12161 Met-core
33	1094.6	87.9	1473	2	AAQ05604	Aaq05604 Streptoki
34	1093.2	87.8	1320	6	ABN70192	Abn70192 Streptoco
35	1050.4	84.4	1068	2	AAX80493	Aax80493 Recombina
36	1045.8	84.0	2208	2	AAX83589	Aax83589 Recombina
37	811.4	65.2	1262	2	AAQ10230	Aaq10230 Synthetic
38	440.4	35.4	450	2	AAX80494	Aax80494 Recombina
c 39	46.2	3.7	2000	7	ADA71938	Ada71938 Rice gene
40	43.8	3.5	2000	7	ADA71938	Ada71938 Rice gene
41	42	3.4	1180	2	AAV84163	Aav84163 Streptoco
42	42	3.4	1181	6	AAV84164	Aav84164 Streptoco
c 43	42	3.4	5504	6	ABL70350	Abl70350 Chemical
c 44	42	3.4	5504	6	AAS61309	Aas61309 Human gen
c 45	42	3.4	6071	6	ABL32367	Abl32367 Human imm

ALIGNMENTS

```
RESULT 1
AAQ20665
ID   AAQ20665 standard; DNA; 1245 BP.
XX
AC   AAQ20665;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   21-APR-1992  (first entry)
XX
DE   SKC-2 streptokinase gene from S.equisimilis.
XX
KW   plasminogen activator; coronary thrombosis; ATCC-9542; ss.
XX
OS   Streptococcus dysgalactiae subsp. equisimilis.
XX
PN   AU9178101-A.
XX
PD   28-NOV-1991.
XX
PF   31-MAY-1991;  91AU-00078101.
XX
PR   23-MAY-1990;  90CU-00000090.
XX
PA   (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.
XX
DR   WPI; 1992-024716/04.
DR   P-PSDB; AAR20202.
XX
PT   Streptokinase C-2 gene from S equisimilis type C - plasmids and
PT   transformants for large scale intra- and extracellular expression of SKC-
PT   2 useful in thrombolytic agents.
XX
PS   Claim 6; Page 14; 28pp; English.
XX
CC   The SKC-2 gene is isolated from S.equisimilis type C by gene
CC   amplification from the synthetic primers SK1, SK2 and SK3 (see AAQ20666-
CC   8). The SKC-2 gene was inserted into yeast expression vector pPS-7 for
CC   extracellular expression of streptokinase and into pNAO for intracellular
CC   expression in yeast. For expression in bacteria, the SKC-2 gene was
CC   inserted in plasmid pEKG3, between a trp promoter and a T4 terminator.
CC   (Updated on 25-MAR-2003 to correct PA field) (Updated on 24-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 1245 BP; 428 A; 273 C; 232 G; 312 T; 0 U; 0 Other;
```

```
Query Match        99.7%;    Score 1241.8;   DB 2;   Length 1245;
Best Local Similarity  99.8%;  Pred. No. 0;
Matches 1243;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy     1 ATTGTGGACCTGAGTCGCTGCTAGACCGTTCCATCTGTCAACAACAGCCAATTAGTTGTT   60
         ||||| |||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db     1 ATTGCTGGACCTGAGTCGGCTGCTAGACCGTTCCATCTGTCAACAACAGCCAATTAGTTGTT   60

Qy    61 AGCGGTTGCTGGTACTCGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AGCGGTTGCTGGTACTCGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT  120

Qy   121 GACCTAACATCAGACGACCTGCTCATGGAGGGAAAGAGACAGAGCAAGGCTTAAGTCCAAAATCA  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GACCTAACATCACGACGACCTGCTCATGGAGGGAAAGAGACAGAGCAAGGCTTAAGTCCAAAATCA  180

Qy   181 AAACCATTTGCTACTGATAGTGGCGGCGATGGCCACATAAAACTGAAAAAGTCGACTTACTA  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAACCATTTGCTACTGATAGTGGCGGCGATGGCCACATAAAACTGAAAAAGTCGACTTACTA  240

Qy   241 AAGGCTATTCAAGAAGAACAATTGATCGCTGCTAACGTCCAGTAACGACGACTACTTTGAGGTC  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AAGGCTATTCAAGAAGAACAATTGATCGCTGCTAACGTCCAGTAACGACGACTACTTTGAGGTC  300

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGACGGCAAGGTCTACTTGCTGAC  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATTGATTTTGCAAGCGATGCAACCATTACTGACGGCAAGGTCTACTTGCTGAC  360

Qy   361 AAAGATGGTTCGGTAACCTTGCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGATGGTTCGGTAACCTTGCGACCCAACCTGTCCAAGAATTTTCCCAAGCGGACAT  420

Qy   421 GTGGCGGCTTAGACCATATAAAGAAAAACAAATCAAGCGAAATCTGTTGATGTG  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTGGCGGCTTAGACCATATAAAGAAAAACAAATCAAGCGAAATCTGTTGATGTG  480

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACCAGGTCTCAAA  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GAATATATCTGTACAGTTTACTCCCTTAAACCCTGATGACCAGGTCTCAAA  540

Qy   541 GATACTAAGCTATTGAAAAACACTAGCTATCGTGACACCATCACATCTCAAGAAATTACTA  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GATACTAAGCTATTGAAAAACACTAGCTATCGTGACACCATCACATCTCAAGAAATTACTA  600

Qy   601 GCTCAAGCACAAAAGCATTTTAAACAAAACCCAGGCTATAACGATTATGAACGTGAC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCTCAAGCACAAAAGCATTTTAAACAAAACCCAGGCTATAACGATTATGAACGTGAC  660

Qy   661 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGACATTTTACCAATGGATCAAGAG  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGACATTTTACCAATGGATCAAGAG  720

Qy   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG  780

Qy   781 AATGAGAGAAATAAACAACAACACTGACCTGATCTCTGAGAAAATATTACGTCCTTAAAAAGGG  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AATGAGAGAAATAAACAACAACACTGACCTGATCTCTGAGAAAATATTACGTCCTTAAAAAGGC  840

Qy   841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAACTGTTCACCATCAAATACGTT  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAACTGTTCACCATCAAATACGTT  900

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC  960

Qy   961 TTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAACTACTTCTACACAATCTC  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAACTACTTCTACACAATCTC  1020

Qy  1021 GATGGCTTTTGGTATTATGGACTATACCTTAACTGGACGTAATCACGATGAC  1080
```

```
Db  1021 GATGGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAAGTAGAGGATAAATCACGATGAC  1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT  1140

Qy  1141 CATTAGCCTATGATAAAGATCGTTATACCGAAGAACGAGAAGTTACAGCTACCTG  1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CATTAGCCTATGATAAAGATCGTTATACCGAAGAACGAGAAGTTACAGCTACCTG  1200

Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA  1245
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA  1245
```

```
RESULT 2
AAQ12156
ID   AAQ12156 standard; DNA; 1335 BP.
XX
AC   AAQ12156;
XX
DT   24-OCT-2003 (revised)
DT   25-MAR-2003 (revised)
DT   17-SEP-1991 (first entry)
XX
DE   Streptokinase gene.
XX
KW   Fusion protein; blood clotting; coagulation; fibrinolysis;
KW   antithrombotic; thrombolysis; streptokinase; ss.
XX
OS   Streptococcus dysgalactiae subsp. equisimilis; ATCC 9542 or ATCC 10009.
XX
FH   Key          Location/Qualifiers
FT   CDS          7..1329
FT                /*tag= a
FT   sig_peptide  7..84
FT                /*tag= b
FT   mat_peptide  85..1326
FT                /*tag= c
XX
PN   WO9109125-A.
XX
PD   27-JUN-1991.
XX
PF   07-DEC-1989;   89GB-00027722.
XX
PR   07-DEC-1989;   89GB-00027722.
PR   07-DEC-1990;   90WO-GB001911.
XX
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
XX
PI   Dawson KM, Hunter MG, Czapleswsk LG;
XX
PW   WPI; 1991-208151/28.
DR   P-PSDB; AAR12889.
XX
PT   Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT   fractions having greater antithrombotic activity for therapy and
PT   prophylaxis.
XX
PS   Disclosure; Page 80; 115pp; English.
XX
CC   The sequence was obtd. from PCR amplified chromosomal DNA from S.
CC   equisimilis (Lancefield's Gp C ) ATCC 10009 or ATCC 9642. The primers
CC   used for the PCR were based on the published DNA sequence of S.
CC   equisimilis strain H46A (Malke, H., Roe, B., and Ferretti, J.J., Gene 34
CC   357-362 (1985)). See Swiss-Prot K02986 and Geneseq N70106. The sequence
CC   is an EcoRI-BamHI fragment. The gene can be used to construct expression
CC   vectors in which the streptokinase gene is linked to a second gene
CC   encoding e.g. another streptokinase protein, hirudin, or a streptokinase-
CC   like protein, via a linking sequence encoding a cleavage site for e.g.
CC   factor Xa or thrombin. The enzymes which cleave the fusion protein are
```

CC   present at the site of the target thrombus so the active agents are
CC   released specific- ally at the place where clot formation is occurring.
CC   See also Q12153-Q12155, Q12158-Q12162 and Q12490. (Updated on 25-MAR-2003
CC   to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 1335 BP; 453 A; 287 C; 250 G; 345 T; 0 U; 0 Other;

Query Match   99.7%; Score 1241.8; DB 2; Length 1335;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1243; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 144

Qy   61 AGCGGTTGCTCGTTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 AGCGGTTGCTCGTTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 204

Qy   121 GACCTAACATCACGACCTGCTCATGGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   205 GACCTAACATCACGACCTGCTCATGGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 264

Qy   181 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTACTA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   265 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTACTA 324

Qy   241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   325 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 384

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGACTGAAAACGGCAAGGTCTACTTTGCTGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   385 ATTGATTTTGCAAGCGATGCAACCATTACTGACTGAAAACGGCAAGGTCTACTTTGCTGAC 444

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGGACAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   445 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGGACAT 504

Qy   421 GTGCGCGTTGAGACCATATAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   505 GTGCGCGTTGAGACCATATAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 564

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACGCAGGTCTCAAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   565 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACGCAGGTCTCAAA 624

Qy   541 GATACTAAGCTATTGAAAACACTAGCTACGTCGGTGACACCATCACATCTCAAGARTTACTA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   625 GATACTAAGCTATTGAAAACACTAGCTACGTCGGTGACACCATCACATCTCAAGARTTACTA 684

Qy   601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   685 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC 744

Qy   661 TCCTCAATCGTCACTCACTCATGACAATTTCCGTACGATTTTCCAATGGATCAAGAG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   745 TCCTCAATCGTCACTCACTCATGACAATTTCCGTACGATTTTCCAATGGATCAAGAG 804

Qy   721 TTTACTTACCATGTCAAAAAATCGGGAACAAGCTTATGATCAATAAAAATCTGGTCTG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   805 TTTACTTACCATGTCAAAAAATCGGGAACAAGCTTATGATCAATAAAAATCTGGTCTG 864

Qy   781 AATGAAGAATAAACAACAACAAATATTACGTCCTTAAAAAGGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   865 AATGAAGAATAAACAACAACAAATATTACGTCCTTAAAAAGGG 924

Qy   841 GAAAAGCCGTATGATCCCTTGATCGCAGTTCACTTGAAAACTGTTCACCATCAAATACGTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   925 GAAAAGCCGTATGATCCCTTGATCGCAGTTCACTTGAAAACTGTTCACCATCAAATACGTT 984

Qy   901 GATGTCAACACCAACGAATTGCTAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   985 GATGTCAACACCAACGAATTGCTAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1044

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1045 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1104

Qy   1021 GATGCTTTTGGTATTATGATGGACTATACTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105 GATGCTTTTGGTATTATGATGGACTATACTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1164

Qy   1081 ACCAACCGATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1224

Qy   1141 CATTTAGCCTATGATAAAGATCGTTAACCGAAGAGAAGATCAGAAGAACGAGAAGTTTACAGCTACCTG 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAAGATCAGAAGAACGAGAAGTTTACAGCTACCTG 1284

Qy   1201 CGTTATACAGGGACACCTATACCTGATAACCCTAAGCACAAATAA 1245
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1285 CGTTATACAGGGACACCTATACCTGATAACCCTAAGCACAAATAA 1329

RESULT 3
AAQ12162
ID   AAQ12162 standard; DNA; 1458 BP.
XX
AC   AAQ12162;
XX
DT   25-MAR-2003 (revised)
DT   17-SEP-1991 (first entry)
XX
DE   Factor Xa-cleavable hirudin-IEGR-streptokinase gene.
XX
KW   Fusion protein; blood clotting; coagulation; fibrinolysis;
KW   antithrombotic; thrombolysis; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1452
FT   /*tag= a
FT   mat_peptide   1..195
FT   /*tag= b
FT   /label= hirudin HV-1
FT   misc_RNA   196..207
FT   /*tag= c
FT   /label= linker
FT   /note= "encodes factor Xa cleavage site"
FT   mat_peptide   208..1449
FT   /*tag= d
FT   /label= streptokinase
XX
PN   WO9109125-A.
XX
PD   27-JUN-1991.
XX
PF   07-DEC-1989; 89GB-00027722.
XX
PR   07-DEC-1989; 89GB-00027722.
PR   07-DEC-1990; 90WO-GB001911.
XX
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
XX
PI   Dawson KM, Hunter MG, Czapleswsk LG;
XX
DR   WPI; 1991-208151/28.
DR   P-PSDB; AAR12885.
XX
PT   Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT   fractions having greater antithrombotic activity for therapy and
PT   prophylaxis.
XX

```
PS   Disclosure; Page 96; 115pp; English.
XX
CC   The sequence of the synthetic hirudin HV-1 genes was designed based on
CC   the published amino acid sequence (Dodt J., et al FEBS Letters 165 180
CC   (1984). The sequence of streptokinase was obtd. from PCR amplified
CC   chromosomal DNA from S. equisimilis ATCC 10009 or ATCC 9642. The primers
CC   used for the PCR were based on the pub- lished DNA sequence of S.
CC   equisimilis strain H46A (Malke, H., Roe, B., and Ferretti, J.J., Gene 34
CC   357-362 [1985]). The two sequences were used to construct an expression
CC   vector in which the hirudin gene is linked to the streptokinase gene via
CC   a linking sequence encoding a cleavage site for factor Xa. The factor Xa
CC   is present at the site of the target thrombus so the active agents are
CC   released specifically at the place where clot formation is occurring. See
CC   also AAQ12153-Q12156, AAQ12158-Q12161 and AAQ12490. (Updated on 25-MAR-
CC   2003 to correct PA field.)
XX
SQ   Sequence 1458 BP; 491 A; 316 C; 290 G; 361 T; 0 U; 0 Other;

Query Match             99.7%;  Score 1241.8;  DB 2;  Length 1458;
Best Local Similarity   99.8%;  Pred. No. 0;
Matches 1243;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   208 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 267

Qy    61 AGCGGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTGTCTTAAATTTTTGAAATT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268 AGCGGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTGTCTTAAATTTTTGAAATT 327

Qy   121 GACCTAACATCACGACCTGCTCATGGAGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 GACCTAACATCACGACCTGCTCATGGAGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 387

Qy   181 AAACCATTTGCTACTGATAGTGGCGGCGATGCCACATAAAACTGAAAAAGTGACTTACTA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388 AAACCATTTGCTACTGATAGTGGCGGCGATGCCACATAAAACTGAAAAAGTGACTTACTA 447

Qy   241 AAGGCTATTCAAGAACAATTGATCGCTCAAGCTCCACAGTCAACGACTACTTGAGGTC 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   448 AAGGCTATTCAAGAACAATTGATCGCTCAAGCTCCACAGTCAACGACTACTTGAGGTC 507

Qy   301 ATTGATTTTGCAAGGCGATGCAACCATTACTGATCGTGATCGAAACGGCAAGGTCTCTACTTTGCTGAC 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508 ATTGATTTTGCAAGGCGATGCAACCATTACTGATCGTGATCGAAACGGCAAGGTCTCTACTTTGCTGAC 567

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568 AAAGATGGTTCGGTAACCTTGCCGACCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 627

Qy   421 GTGCGCGGTTAGACCATTAGAAGAAAACAAAATCAAGCGAAATCTGTTGATGTG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   628 GTGCGCGGTTAGACCATTAGAAGAAAACAAAATCAAGCGAAATCTGTTGATGTG 687

Qy   481 GAATAACTGTACAGTTTACTCCCTTAAACCCGATGACGAGTTTCAGACCAGGTCTCAAA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   688 GAATAACTGTACAGTTTACTCCCTTAAACCCGATGACGAGTTTCAGACCAGGTCTCAAA 747

Qy   541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTCGTACCACCCACCAGCTATACGTGAC 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   748 GATACTAAGCTATTGAAAAACACTAGCTATCGGTCGTACCACCCACCAGCTATACGTGAC 807

Qy   601 GCTCAAGCACAAAGCATTTAAACAAAACCCAGCCTATACGATTATGAACGTGAC 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   808 GCTCAAGCACAAAGCATTTAAACAAAACCCAGCCTATACGATTATGAACGTGAC 867

Qy   661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   868 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 927

Qy   721 TTTACTTACCATGTCAAAAATGGGAACAAGCTTATGAGATCAATAAAAAATCTCGTCTG 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   928 TTTACTTACCATGTCAAAAATGGGAACAAGCTTATGAGATCAATAAAAAATCTCGTCTG 987

Qy   781 AATGAGAAATAAACACACTGACCTGATCCTGAGAAATATTACGTCCTTAAAAAAGGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   988 AATGAGAAATAAACACACTGACCTGATCCTGAGAAATATTACGTCCTTAAAAAAGGG 1047

Qy   841 GAAAAGCCGTATGATCCTTTGATCGGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1048 GAAAAGCCGTATGATCCTTTGATCGGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 1107

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1108 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1167

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTTCTACAACAATCTC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1168 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTTCTACAACAATCTC 1227

Qy  1021 GATGCTTTTGGTATTATGGGACTATACCTTAACTGGATACATACCTGGAAAAGTAGAAGATAATCACGATGAC 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1228 GATGCTTTTGGTATTATGGGACTATACCTTAACTGGATACATACCTGGAAAAGTAGAAGATAATCACGATGAC 1287

Qy  1081 ACCAACCGTATCATAACGCTTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1288 ACCAACCGTATCATAACGCTTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1347

Qy  1141 CATTAGGCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1348 CATTAGGCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1407

Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTATAACCGATAACCTAACGACAAATAA 1245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1408 CGTTATACAGGGACACCTATACCTGATAACCCTATAACCGATAACCTAACGACAAATAA 1452


RESULT 4
AAQ12158
ID   AAQ12158 standard; DNA; 1512 BP.
XX
AC   AAQ12158;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   17-SEP-1991  (first entry)
XX
DE   Streptokinase gene fused to yeast alpha factor secretion sequence.
XX
KW   Fusion protein; blood clotting; coagulation; fibrinolysis;
KW   antithrombotic; thrombolysis; streptokinase; ss.
XX
OS   Streptococcus dysgalactiae subsp. equisimilis; ATCC 9542 or ATCC 10009.
XX
FH   Key             Location/Qualifiers
FT   CDS             7..1506
FT                   /*tag= a
FT   sig_peptide     7..261
FT                   /*tag= b
FT                   /label= pre-pro alpha factor secretion leader
FT   mat_peptide     262..1503
FT                   /*tag= c
FT                   /label= mature streptokinase
XX
PN   WO9109125-A.
XX
PD   27-JUN-1991.
XX
PF   07-DEC-1989;   89GB-00027722.
XX
PR   07-DEC-1989;   89GB-00027722.
PR   07-DEC-1990;   90WO-GB001911.
XX
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
XX
PI   Dawson KM, Hunter MG, Czapleswbk LG;
```

```
Db    922 TCCTCAATGTCACTCATGACAATGACACATTTTCCGTACGATTTTACCAATGGATCAAGAG  981
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    721 TTTACTTACCATGTCAAAAATCGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  780
              ||   |||||||||||||||||||||||||||||||||||||||||||||||
Db    982 TTTACTTACCATGTCAAAAATCGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  1041
Qy    781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCTTAAAAAAGGG  840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1042 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCTTAAAAAAGGG  1101
Qy    841 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  1161
Qy    901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCCTAACAGCTAGCGAACGTAAC  960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1162 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCCTAACAGCTAGCGAACGTAAC  1221
Qy    961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC  1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1222 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC  1281
Qy   1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGAAAAGTAGAGGATAATCACGATGAC  1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1282 GATGCTTTTGGTATTATGGACTATACCTTAACTGAAAAGTAGAGGATAATCACGATGAC  1341
Qy   1081 ACCAACCGTATCATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAAGATGCTAGCTAT  1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1342 ACCAACCGTATCATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAAGATGCTAGCTAT  1401
Qy   1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1402 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  1461
Qy   1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA  1245
              |||||||||||||||||||||||||||||||||||||||||||
Db   1462 CGTTATACAGGGACACCTATACCCTGATAACCCTAACGACACAAATAA  1506
```

```
RESULT 5
AAQ12160
ID   AAQ12160 standard; DNA; 2589 BP.
XX
AC   AAQ12160;
XX
DT   24-OCT-2003   (revised)
DT   25-MAR-2003   (revised)
DT   17-SEP-1991   (first entry)
XX
DE   OmpAL streptokinase-streptokinase gene.
XX
KW   Fusion protein; blood clotting; coagulation; fibrinolysis;
KW   antithrombotic; thrombolysis; major outer membrane protein A; thrombin;
KW   ss.
XX
OS   Streptococcus dysgalactiae subsp. equisimilis; ATCC 9542 or ATCC 10009.
XX
FH   Key             Location/Qualifiers
FT   CDS             4..2583
FT                   /*tag= a
FT   sig_peptide     4..66
FT                   /*tag= b
FT                   /label= ompAL signal sequence
FT                   /note=" E.coli"
FT   mat_peptide     67..1308
FT                   /*tag= c
FT                   /label= mature streptokinase
FT   mat_peptide     1339..2580
FT                   /*tag= e
FT                   /label= mature streptokinase
FT   CDS             1809..1338
FT                   /*tag= d
FT                   /label= linker
```

```
XX
DR   WPI; 1991-208151/28.
DR   P-PSDB; AAR12891.
XX
PT   Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT   fractions having greater antithrombotic activity for therapy and
PT   prophylaxis.
XX
PS   Disclosure; Page 86; 115pp; English.
XX
CC   The streptokinase sequence was obtd. from PCR amplified chromosomal DNA
CC   from S. equisimilis (Lancefield's Gp C) ATCC 10009 or ATCC 9642 (the
CC   primers used for the PCR were based on the published DNA sequence of S.
CC   equisimilis strain H46A (Malke, H., Roe, B., and Ferretti, J.J., Gene 34
CC   357-362 [1985]). The gene was fused to DNA encoding the yeast alpha
CC   factor pre-pro-secretion sequence in an expression vector, pSMD1/152, for
CC   prodn. of recombinant strepto- kinase in S. cerevisiae strain BJ2168. See
CC   also AAQ12153-Q12156, AAQ12159-Q12162 and AAQ12490. (Updated on 25-MAR-
CC   2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC   field)
XX
SQ   Sequence 1512 BP; 508 A; 323 C; 288 G; 393 T; 0 U; 0 Other;

Query Match             99.7%;  Score 1241.8;  DB 2;  Length 1512;
Best Local Similarity   99.8%;  Pred. No. 0;
Matches 1243;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;
```

```
Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTAGTTGTT  60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTAGTTGTT  321
Qy    61 AGCGTTGCTGGTACTGTTGAGGGGACGAACATTAGTCTTAAATTTTTGAAATT  120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322 AGCGTTGCTGGTACTGTTGAGGGGACGAACATTAGTCTTAAATTTTTGAAATT  381
Qy   121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   382 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  441
Qy   181 AAACCATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAAAAGCTGACTTACTA  240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   442 AAACCATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAAAAGCTGACTTACTA  501
Qy   241 AAGGCTATTCAAGAACAATTGATCGTCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   502 AAGGCTATTCAAGAACAATTGATCGTCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  561
Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGACTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC  360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   562 ATTGATTTTGCAAGCGATGCAACCATTACTGACTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC  621
Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT  420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   622 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT  681
Qy   421 GTGCGCGTTAGACCATATAAAGAAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG  480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   682 GTGCGCGTTAGACCATATAAAGAAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG  741
Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   742 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  801
Qy   541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCCATCTCAAGAATTACTA  600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   802 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCCATCTCAAGAATTACTA  861
Qy   601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGCTATACGATTTATGAACGTGAC  660
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   862 GCTCAAGCACAAAGCATTTAAACAAAACCCATCCAGGCTATACGATTTATGAACGTGAC  921
Qy   661 TCCTCAATGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG  720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
FT                       /note= "encodes thrombin cleavage site"
XX
PN   WO9109125-A.
XX
PD   27-JUN-1991.
XX
PF   07-DEC-1989;  89GB-00027722.
XX
PR   07-DEC-1989;  89GB-00027722.
PR   07-DEC-1990;  90WO-GB001911.
XX
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
XX
PI   Dawson KM, Hunter MG, Czapleswsk LG;
XX
DR   WPI; 1991-208151/28.
DR   P-PSDB; AAR12893.
XX
PT   Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT   fractions having greater antithrombotic activity for therapy and
PT   prophylaxis.
XX
PS   Disclosure; Page 90; 115pp; English.
XX
CC   The streptokinase sequence was obtd. from PCR amplified chromosomal DNA
CC   from S. equisimilis ATCC 10009 or ATCC 9642. The primers used for the PCR
CC   were based on the published DNA sequence of S. equisim- ilis strain H46A
CC   (Malke, H., Roe, B., and Ferretti, J.J., Gene 34 357-362 (1985)). The
CC   gene was used to construct an expression vector in which the two
CC   streptokinase genes are linked via a link- ing sequence. This encodes a
CC   cleavage site for thrombin which, when cleaved, releases the individual
CC   proteins which have anti- thrombotic activity. The thrombin is present at
CC   the site of the target thrombus so the active agents are released
CC   specifically at the place where clot formation is occurring. See also
CC   AAQ12153Q12156, AAQ12158-Q12162 and AAQ12490. (Updated on 25-MAR-2003 to
CC   correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 2589 BP; 879 A; 564 C; 500 G; 646 T; 0 U; 0 Other;

Query Match             99.7%;  Score 1241.8;  DB 2;  Length 2589;
Best Local Similarity   99.8%;  Pred. No. 0;
Matches 1243;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 1398

Qy     61 AGCGGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 AGCGGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 1458

Qy    121 GACCTAACATCACGACACCTGCTCATGGAGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 GACCTAACATCACGACACCTGCTCATGGAGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 1518

Qy    181 AAACCATTTGCTACTGATAGTGGCGCGGATGCGCACATAAACTTGAAAAAGCTGACTTACTA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519 AAACCATTTGCTACTGATAGTGGCGCGGATGCGCACATAAACTTGAAAAAGCTGACTTACTA 1578

Qy    241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 1638

Qy    301 ATTGATTTTGCAAGCGATGCAACCATTACTGATGCAACCATTACTGATGCAACCAACGTCTGCTAAAATTTTTGCTGAC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 ATTGATTTTGCAAGCGATGCAACCATTACTGATGCAACCATTACTGATGCAACCAACGGCTCTGCTAAAATTTTTGCTGAC 1698

Qy    361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 1758

Qy    421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 1818

Qy    481 GAATATACTGTACAGTTACTCCCTTAAACCTGATGACGATTTCAGACCAGGTCTCAAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA 1878

Qy    541 GATACTAAGCTATTGAAAACACTAGCTAGTCATCGGTGACACCATCACATCTCAAGAATTACTA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 GATACTAAGCTATTGAAAACACTAGCTAGTCATCGGTGACACCATCACATCTCAAGAATTACTA 1938

Qy    601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 GCTCAAGCACACAAAGCATTTTAAACAAAACCATCCCAGGCTATACGATTTATGAACGTGAC 1998

Qy    661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 2058

Qy    721 TTTACTTACCATGTCAAAAATCGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 TTTACTTACCATGTCAAAAATCGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 2118

Qy    781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 2178

Qy    841 GAAAAGCCGTATGATCCCTTTGATCGGCAGTCACTTGACTTGAAACTGTTCACCACAAATACGTT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2179 GAAAAGCCGTATGATCCCTTTGATCGGCAGTCACTTGACTTGAAACTGTTCACCATCAAATACGTT 2238

Qy    901 GATGTCAACACCAACCAAGAATTGCTAAAAAGCAGCAGCTCTTAACAGCTAGCGAACGTAAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2239 GATGTCAACACCAACCAACGAATTGCTAAAAAGCAGCAGCTCTTAACAGCTAGCGAACGTAAC 2298

Qy    961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGTGAAAAGTAGAGGATAATCACGATGAC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2299 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGTGAAAAGTAGAGGATAATCACGATGAC 2358

Qy   1021 GATGCTTTTGGTATTATGGACTATCATACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGTAT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2359 GATGCTTTTGGTATTATGGACTATACCGTTTATATGGCAAGCGACCGAAGGAGAGAATGCTAGTAT 2418

Qy   1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCGCAAGGCGACCGAAGGAGAGAAGAATGCTAGTAT 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2419 ACCAACCGTATCATAACCGTTTATATGGCAAGCGACCGCAAGGCGACCGAAGGAGAGAAGAATGCTAGTAT 2478

Qy   1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGATCGTTATACCGAAGAGAACAGAGAAGTTTACAGCTACCTG 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2479 CATTTAGCCTATGATAAAGATCGTTATACCGAAGATCGTTATACCGAAGAGAACAGAGAAGTTTACAGCTACCTG 2538

Qy   1201 CGTTATACAGGGACACCTACCTGATAACCCTACCTGATAACCCTACCTAACGACAAATAA 1245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2539 CGTTATACAGGGACACCTACCTGATAACCCTACCTGATAACCCTACCTAACGACAAATAA 2583
```

```
RESULT 6
AAQ12490
ID   AAQ12490 standard; DNA; 1467 BP.
XX
AC   AAQ12490;
XX
DT   25-MAR-2003  (revised)
DT   17-SEP-1991  (first entry)
XX
DE   Factor Xa-cleavable streptokinase-IEGR-hirudin gene.
XX
KW   Fusion protein; blood clotting; coagulation; fibrinolysis;
KW   antithrombotic; thrombolysis; ss.
XX
OS   Synthetic.
XX
FH   Key                 Location/Qualifiers
FT   CDS                 1..1455
```

```
FT   mat_peptide      /*tag= a
FT                    1..1242
FT                    /*tag= b
FT                    /label= streptokinase
FT   misc_RNA         1243..1254
FT                    /*tag= c
FT                    /label= linker
FT                    /note= "encodes factor Xa cleavage site"
FT   mat_peptide      1255..1453
FT                    /*tag= d
FT                    /label= hirudin HV-1
XX
PN   WO9109125-A.
XX
PD   27-JUN-1991.
XX
PF   07-DEC-1989;     89GB-00027722.
XX
PR   07-DEC-1989;     89GB-00027722.
PR   07-DEC-1990;     90WO-GB001911.
XX
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
XX
PI   Dawson KM,  Hunter MG,  Czapleswk LG;
XX
DR   WPI; 1991-208151/28.
DR   P-PSDB; AAR12522.
XX
PT   Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT   fractions having greater antithrombotic activity for therapy and
PT   prophylaxis.
XX
PS   Disclosure; Page 98; 115pp; English.
XX
CC   The sequence of the synthetic hirudin HV-1 gene was designed based on the
CC   published amino acid sequence (Dodt J., et al FEBS Letters 165 180
CC   (1984)). The sequence of streptokinase was obtd. from PCR amplified
CC   chromosomal DNA from S. equisimilis ATCC 10009 or ATCC 9642. The primers
CC   used for the PCR were based on the pub-lished DNA sequence of S.
CC   equisimilis strain H46A (Malke, H., Roe, B., and Ferretti, J.J., Gene 34
CC   357-362 (1985)). The two sequences were used to construct an expression
CC   vector in which the streptokinase gene is linked to the hirudin gene via
CC   a linking sequence encoding a cleavage site for factor Xa. The factor Xa
CC   is present at the site of the target thrombus so the active agents are
CC   released specifically at the place where clot formation is occurring. See
CC   also AAQ12153-Q12156 and AAQ12158-Q12162. (Updated on 25-MAR-2003 to
CC   correct PA field.)
XX
SQ   Sequence 1467 BP; 494 A; 317 C; 292 G; 364 T; 0 U; 0 Other;

Query Match             99.5%;  Score 1238.8;  DB 2;  Length 1467;
Best Local Similarity   99.8%;
Matches 1240;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    0;

Qy        1 ATTGCTGGACCTGAGTGGGCTGCTAGACCGTTCCATCGTCAACAACAGCCAATTAGTTGTT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATTGCTGGACCTGAGTGGGCTGCTAGACCGTTCCATCGTCAACAACAGCCAATTAGTTGTT   60

Qy       61 AGGGTTGCTGGTACTGTTGAGGGGACGAATCAAGAACATTAGTCTTAAATTTTTTGAAATT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 AGGGTTGCTGGTACTGTTGAGGGGACGAATCAAGAACATTAGTCTTAAATTTTTTGAAATT  120

Qy      121 GACCTAACATCACGACCTGCTCATGGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA   180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GACCTAACATCACGACCTGCTCATGGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA   180

Qy      181 AAACCATTTGCTACTGATAGTGGCGCGATGGCCACATAAACTTGAAAAAGCTGACTTACTA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 AAACCATTTGCTACTGATAGTGGCGCGATGGCCACATAAACTTGAAAAAGCTGACTTACTA  240

Qy      241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300

Qy      301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGTCTACTTTGCTGAC   360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGTCTACTTTGCTGAC   360

Qy      361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT   420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT   420

Qy      421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG   480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG   480

Qy      481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACACCAGGTCTCAA   540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACACCAGGTCTCAA   540

Qy      541 GATACTAAGCTATTGAAAACACTAGCTATCGTGACACCATCACATCTCAAGAATTACTA   600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GATACTAAGCTATTGAAAACACTAGCTATCGTGACACCATCACATCTCAAGAATTACTA   600

Qy      601 GCTCAAGCACAAAGCATTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC   660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 GCTCAAGCACAAAGCATTTAAACAAAACCCATCCAGGCTATACGATTTATGAACGTGAC   660

Qy      661 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG   720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG   720

Qy      721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG   780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG   780

Qy      781 AATGAAGAAATAAACAACACTGACCTGACTCTCTGAGAAATATTAGTCCTAAAAAGGG   840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 AATGAAGAAATAAACAACACTGACCTGACTCTCTGAGAAATATTAGTCCTAAAAAGGG   840

Qy      841 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTGTTCACCATCAAATACGTT   900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTGTTCACCATCAAATACGTT   900

Qy      901 GATGTCAACACCAAGCGAATTGCTAAAAAGGCAGCTCTTAACAGCTAGGGAACGTAAC   960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 GATGTCAACACCAAGCGAATTGCTAAAAAGGCAGCTCTTAACAGCTAGGGAACGTAAC   960

Qy      961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGCTAAACTACTCTACAACAATCTC  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGCTAAACTACTCTACAACAATCTC  1020

Qy     1021 GATGCTTTCGTATTCGGTATTATGGACTATACCTTAACTGAAAAGTAGAGGATAATCACGATGAC  1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 GATGCTTTCGTATTCGGTATTATGGACTATACCTTAACTGAAAAGTAGAGGATAATCACGATGAC  1080

Qy     1081 ACCAACGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT  1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 ACCAACGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT  1140

Qy     1141 CATTTAGCCTATGATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 CATTTAGCCTATGATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  1200

Qy     1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA  1242
            ||||||||||||||||||||||||||||||||||||||||||
Db     1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA  1242

RESULT 7
AAA37633
ID   AAA37633 standard; DNA; 1245 BP.
XX
AC   AAA37633;
XX
DT   15-SEP-2003 (revised)
```

DT   13-OCT-2000 (first entry)
XX
DE   S. equisimilis streptokinase coding sequence.
XX
KW   Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW   plasminogen; human; fibronectin; thrombolytic therapy;
KW   cardiovascular disorder; fibronectin; ss.
XX
OS   Streptococcus dysgalactiae subsp. equisimilis.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1245
FT   /*tag= a
FT   /product= "streptokinase"
XX
PN   EP1024192-A2.
XX
PD   02-AUG-2000.
XX
PF   23-DEC-1999; 99EP-00310541.
XX
PR   24-DEC-1998; 98IN-DE003825.
XX
PA   (COUL) CSIR COUNCIL SCI IND RES.
XX
PI   Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI   Yadav M;
XX
DR   WPI; 2000-516032/47.
DR   P-PSDB; AAY90282.
XX
PT   Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT   thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT   domains of human fibronectin.
XX
PS   Example 3; Fig 3; 58pp; English.
XX
CC   This sequence represents the human Streptococcus equisimilus
CC   streptokinase coding sequence. The invention relates to a hybrid
CC   plasminogen activator (PA) comprises a polypeptide fusion between
CC   streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC   fibrin binding regions of human fibronectin, which are from fibrin
CC   binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC   ability to bind with fibrin independently and also characteristically
CC   retains a PG activation ability which becomes evident only after a
CC   pronounced duration, or lag, after exposure of the PA to a suitable
CC   animal or human PG. The hybrid streptokinase-fibrin binding domain
CC   polypeptides are useful in thrombolytic therapy for various kinds of
CC   cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC   well as kinetics of plasminogen activation that are distinct from that of
CC   natural streptokinase being characterised by a temporary delay, or lag
CC   of several minutes in the natural rate of the catalytic conversion of
CC   plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC   can bind tightly with fibrin in blood clots soon after introduction into
CC   the vascular system without significantly activating the circulating
CC   blood plasminogen to plasmin, thus aiding in the localisation of the
CC   plasminogen activation process to the site of pathological thrombus. This
CC   overcomes systemic plasminogen activation encountered during clinical use
CC   of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ   Sequence 1245 BP; 426 A; 267 C; 237 G; 315 T; 0 U; 0 Other;

Query Match   98.3%; Score 1224.2; DB 3; Length 1245;
Best Local Similarity   99.0%; Pred. No. 0;
Matches 1232;   Conservative   0; Mismatches   13; Indels   0; Gaps   0;

Qy   1 ATTGCTCGACCTGAGTCGGCTGCTGAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATTGCTCGACCTGAGTCGGCTGCTGAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT 60

Qy   61 AGCGGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTTGAAATT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGCGGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTTGAAATC 120

Qy 121 GACCTAACATCACGACCTGCTCATGGAGGAAGAACAGAGCAAGGCTTAAGTCCAAAATCA 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GATCTAACATCACGACCTGCTCATGGAGGAAGAACAGAGCAAGGCAAGGCTTAAGTCCAAAATCA 180

Qy 181 AAACCATTTGCTACTGATAGTGGCGGCGATGCCGCGATGTCACATAAACTGAGAAAGGCTGACTTACTA 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 AAACCATTTGCTACTGATAGTGGCGGCGATGCCGCGATGTCACATAAACTGAGAAAGGCTGACTTACTA 240

Qy 241 AAGGCTATTCAAGAACAATTGATCGTCAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 AAGGCTATTCAAGAACAATTGATCGTCAACGTCCACAGTAACGACGACTACTTTGAGGTC 300

Qy 301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGCAACCATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGCAACCATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 360

Qy 361 AAAGATGGTTCGGTAACCTTGCCGACCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420

Qy 421 GTGCGCGGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTCGTTGATGTG 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GTGCGCGGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTCGTTGATGTG 480

Qy 481 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACCACCAGTCTCAAA 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACCACCAGTCTCAAA 540

Qy 541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACATCACATCTCAAGAATTACTA 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCACATCTCAAGAATTACTA 600

Qy 601 GCTCAAGCACAACCTTTAAACAAAACCACCAGGCTATACGATTTATGAACGTGAC 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 GCTCAAGCACAACCTTTAAACAAAACCACCAGGCTATACGATTTATGAACGTGAC 660

Qy 661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCATGATCAAGAG 720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCATGATCAAGAG 720

Qy 721 TTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 TTTACTTACCGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAATCTGGTCTG 780

Qy 781 AATGAAGAAATAAACACACTGACCTCTGAGAAATATTACGGTCCTTAAAAAAGGG 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 AATGAAGAAATAAACACACTGACCTGACTCTGAGAAATATTACGGTCCTTAAAAAAGGG 840

Qy 841 GAAAAGCCGTATGATCCCTTTGATCGGAGTCACTGGAACTGTTCACCATCAAATACGTT 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 GAAAAGCCGTATGATCCCTTTGATCGGAGTCACTGGAACTGTTCACCATCAAATACGTT 900

Qy 901 GATGTCAACACCAACGAATTGCTAAAAAGCGGAGCAGCTCTTAACAGTCTAGGCAACGTAAC 960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 GATGTCGATACCAACGAATTGCTAAAAAGTGCTAAAAAGTGAGCAGCTCTTAACAGTCTAGGCAACGTAAC 960

Qy 961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020

Qy 1021 GATGCTTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 GATGCTTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1080

Qy 1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACACCCGAAGCGAGAGGAAGAATGCTAGCTAT 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACACCCGAAGCGAGAGGAAGAATGCTAGCTAT 1140

Qy 1141 CATTTAGCCTATGATAAAGATCGTTATACCGGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CATTTAGCCTATGATAAAGATCGTTATACCGGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200

```
Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
```

RESULT 8
AAQ05603
ID   AAQ05603 standard; DNA; 1473 BP.
XX
AC   AAQ05603;
XX
DT   25-MAR-2003 (revised)
DT   20-DEC-1990 (first entry)
XX
DE   Streptokinase G gene.
XX
KW   Streptokinase G; thrombolytic agent; plasminogen; plasmin;
KW   recombinant plasmid; ss.
XX
OS   Streptococcus pyogenes.
XX
FH   Key   Location/Qualifiers
FT   CDS   60..1380
FT   /*tag= a
FT   /product= "streptokinase G"
XX
PN   DD276693-A.
XX
PD   07-MAR-1990.
XX
PF   07-NOV-1988; 88DD-00321531.
XX
PR   07-NOV-1988; 88DD-00321531.
XX
PA   (DEAK) AKAD WISSENSCHAFTEN DDR.
XX
PI   Walter F, Siegel M, Malke H;
XX
DR   WPI; 1990-247327/33.
DR   P-PSDB; AAR06377.
XX
PT   High yield streptokinase G prodn. from recombinant cells - transformed
PT   with plasmid contg. gene from Streptococcus, useful as thrombolytic
PT   agent.
XX
PS   Disclosure; Fig 2; 7pp; German.
XX
CC   Plasmid pMW1 (contg. the streptokinase G gene) is ligated with pUC19
CC   and the resultant product used to transform bacteria. Infected cells are
CC   cultured in lig. medium contg. assimilable C and N sources, and
CC   streptokinase G is isolated from the cell lysate. The infected cells
CC   provide the protein in high yields, i.e. 600 U/ml which is 1.5-2 times
CC   that for the donor strain. See also DD-276694 (AAQ05604). (Updated on 25-
CC   MAR-2003 to correct PA field.)
XX
SQ   Sequence 1473 BP; 498 A; 307 C; 272 G; 396 T; 0 U; 0 Other;

Query Match   98.3%; Score 1224.2; DB 2; Length 1473;
Best Local Similarity   99.0%; Pred. No. 0;
Matches 1232; Conservative   0; Mismatches   13; Indels   0; Gaps   0;
```

```
Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   139 ATTGCTGGACCTGAGTGGCTGCTAGGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 198

Qy    61 AGCGTTGCTGCTGTGTTGAGGGGACGGACGAATCAAGACATTAGTTCTTAAATTTTTGAAATT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   199 AGCGTTGCTGCTGTGTTGAGGGGACGGACGAATCAAGACATTAGTTCTTAAATTTTTGAAATT 258

Qy   121 GACCTAACATCACGACACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCAAAAATCA 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   259 GACCAACATCACGACACCTGCTCACGGAGCAGGGCTTAAGTCAAAAATCA 318
```

```
Qy   181 AAACCATTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTACTA 240
         |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   319 AAACTATTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTACTA 378

Qy   241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   379 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 438

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGATCGAAACGCAAGGTCTACTTTGCTGAC 360
         ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   439 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGATCGAAACGGCAAGGTCTACTTTGCTGAC 498

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCACCGTCCAAGAATTTTGCTCAAGCGGACAT 420
         ||||||||||||||||||||||||| ||||||||||||||||||| || |||||||||
Db   499 AAAGATGGTTCGGTAACCTTGCCGATCCACCGTCCAAGAATTTTGTGTTAAAGGGACAT 558

Qy   421 GTGCGCGTTAGACCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAAATCTGTTGATGTG 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   559 GTGCGCGTTAGACCGTTAGACCATATAAAGAAAAACCAGTACAAAATCAAGCGAAATCTGTTGATGTG 618

Qy   481 GAAATACTAGTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACGACCAGGTCTCAAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   619 GAAATACTAGTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGCTCTCAAA 678

Qy   541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   679 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 738

Qy   601 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTATGAACGTGAC 660
         ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   739 GCTCAAGCACAAAGCATTTTAAACAAAACCACCCAGGCTATACGATTATGAACGTGAC 798

Qy   661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   799 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 858

Qy   721 TTTACTTACCATGTCAAAAATCGGAACAAGCTTATGAGATCAATAAAAAATCTGTCTG 780
         |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   859 TTTACTTACCATGTCAAAAATCGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 918

Qy   781 AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   919 AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 978

Qy   841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   979 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 1038

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGGAGCAGCTCTTAACAGCAGGAACGTAAC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1039 GATGTCAACACCAACGAATTGCTAAAAAGGAGCAGCTCTTAACAGCTAGGGAACGTAAC 1098

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAATCTTACTACTACAACATCTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAATCTCTACTACTCAACAATCTC 1158

Qy  1021 GATGCTTTTGGTATTATGATGGACTATACCTTAACTGGCAAAAGTAGAGGATAATCACGATGAC 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159 GATGCTTTTGGTATTATGATGGACTATACCTTAACTGGCAAAAGTAGAGGATAATCACGATGAC 1218

Qy  1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGGAAGCGAGGAGAATGCTAGCTAT 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGGAAGCGAGGAGAATGCTAGCTAT 1278

Qy  1141 CATTTAGCCTATGATGAAAGATCGTATAAAGATCGTATACCGAAGAGAACCAGAAGTTACAGCTACCTG 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279 CATTTAGCCTATGATGAAAGATCGTATACCGAAGAGAACGAGAAGTTTACAGCTACCTG 1338

Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1339 CGTTATACAGGGACACCTATACCTGATAACCCTAAGTAGTCCAAAATCA 1383
```

```
RESULT 9
AAA37622
ID   AAA37622 standard; DNA; 1377 BP.
XX
AC   AAA37622;
XX
DE   Streptokinase-NTRN gene.
XX
KW   Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW   plasminogen; human; fibronectin; thrombolytic therapy;
KW   cardiovascular disorder; ss.
XX
OS   Streptococcus dysgalactiae subsp. equisimilis.
XX
PN   EP1024192-A2.
XX
PD   02-AUG-2000.
XX
PF   23-DEC-1999;   99EP-00310541.
PR   24-DEC-1998;   98IN-DE003825.
XX
PA   (COUL ) CSIR COUNCIL SCI IND RES.
XX
PI   Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI   Yadav M;
XX
DR   WPI; 2000-516032/47.
XX
PT   Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT   thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT   domains of human fibronectin.
XX
PS   Example 1; Fig 11; 58pp; English.
XX
CC   This sequence represents a streptokinase-NTRN (SK-NTRN) gene (where NTRN
CC   stands for N-terminally repaired with native sequence). The invention
CC   relates to a hybrid plasminogen activator (PA) comprises a polypeptide
CC   fusion between streptokinase (SK), which are capable of plasminogen (PG)
CC   activation, and fibrin binding regions of human fibronectin, which are
CC   from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
CC   possesses the ability to bind with fibrin independently and also
CC   characteristic retains a PG activation ability which becomes evident
CC   only after a pronounced duration, or lag, after exposure of the PA to a
CC   suitable animal or human PG. The hybrid streptokinase-fibrin binding
CC   domain polypeptides are useful in thrombolytic therapy for various kinds
CC   of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
CC   as well as kinetics of plasminogen activation that are distinct from that
CC   of natural streptokinase in being characterised by a temporary delay, or
CC   lag of several minutes in the natural rate of the catalytic conversion of
CC   plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC   can bind tightly with fibrin in blood clots soon after introduction into
CC   the vascular system without significantly activating the circulating
CC   blood plasminogen to plasmin, thus aiding in the localisation of the
CC   plasminogen activation process to the site of pathological thrombus. This
CC   overcomes systemic plasminogen activation encountered during clinical use
CC   of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ   Sequence 1377 BP; 465 A; 302 C; 265 G; 345 T; 0 U; 0 Other;

Query Match      98.2%;  Score 1222.6;  DB 3;  Length 1377;
Best Local Similarity   98.9%;
Matches 1231;  Conservative   0;  Mismatches  14;  Indels   0;  Gaps   0;

Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCGTGTCAACAACAGCCAATTAGTTGTT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   133 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCGTGTCAACAACAGCCAATTGGTTGTT 192

Qy    61 AGCGTTGCTCGGTACTGTTGAGGGGACCGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
          ||||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||
Db   193 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC 252

Qy   121 GACCTAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253 GATCTAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 312

Qy   181 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAAGCTGAAAAAGCTGACTTACTA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   313 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAAGCTGAAAAAGCTGACTTACTA 372

Qy   241 AAGGCTATTCAAGAGACAATGATCGGCTAACGTCCACAGTAACGACGACTACTTGAGGTC 300
          |||||||||||||| ||||| |||||||||||||||||||||||||||||||| |||||
Db   373 AAGGCTATTCAAGAACAATTGATCGGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 432

Qy   301 ATTGATTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTGCTGAC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   433 ATTGATTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 492

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   493 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 552

Qy   421 GTGCGCGTTAGACCATATAAAGAAAAAACCAATACAAAATCAAGCGAAAATCTGTTGATGTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   553 GTGCGCGTTAGACCATATAAAGAAAAAACCAATACAAAATCAAGCGAAAATCTGTTGATGTG 612

Qy   481 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGTCTCAAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   613 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGTCTCAAA 672

Qy   541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 600
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   673 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 732

Qy   601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC 660
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   733 GCTCAAGCACAAAGCATTTAAACAAAAACCACCCAGGCTATACGATTTATGAACGTGAC 792

Qy   661 TCCTCAATGCTCACTCATGACAATGACAATTTCCGTACGATTTACCAATGGATCAAGAG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   793 TCCTCAATGCTCACTCATGACAATGACAATTTCCGTACGATTTACCAATGGATCAAGAG 852

Qy   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
          ||||||||||| || ||||||||||||||||||||| ||||||||||||||||||||||
Db   853 TTTACTTACCGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAAATCTGGTCTG 912

Qy   781 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
          |||||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||
Db   913 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAAATATTACGTCCTTAAAAAAGGG 972

Qy   841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   973 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 1032

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 960
          ||||| ||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  1033 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1092

Qy   961 TTAGACTTCAGAGATTTATACGATCCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1093 TTAGACTTCAGAGATTTATACGATCCGTGATAAGGCTAAACTACTCTACAACAATCTC 1152

Qy  1021 GATGCTTTGGTATTGATGACTATACCTTAACTGGAGGATAATCACGATGAC 1080
          ||||||| |||||||||||||||||||||||||||||||||||||||||||
Db  1153 GATGCTTTTGGTATTGATGACTATACCTTAACTGGAGGATAATCACGATGAC 1212

Qy  1081 ACCAACCGTATCATAACCGTTTATATGGAAGCGACCCGGAAGGAGAAGAATGCTAGCTAT 1140
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  1213 ACCAACCGTATCATAACCGTTTATATGGGAAGCGACCCGGAAGGAGAAGAATGCTAGCTAT 1272

Qy  1141 CATTTAGCCTATGATAAAGATCGTTATCCGAAGAAGAACGAGAAGAAGTTTACAGCTACCTG 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1273 CATTTAGCCTATGATAAAGATCGTTATCCGAAGAAGAACGAGAAGAAGTTTACAGCTACCTG 1332
```

```
Matches  1231;  Conservative   0;  Mismatches   14;  Indels   0;  Gaps   0;

Qy      1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTGGTTGTT   597

Qy     61 AGCGTTGCTGGTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 AGCGTTGCTGGTGCTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC   657

Qy    121 GACCTAACATCACGACCTGCTCATGGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    658 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA   717

Qy    181 AAACCATTTGCTACTGATAGTGGCGCAGTGCCACATAAACTGAAAAAGCTGACTTACTA   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    718 AAACCATTTGCTACTGATAGTGGCGCAGTGCCACATAAACTGAGAAAAGCTGACTTACTA   777

Qy    241 AAGGCTATTCAAGAACAATTGATCGTAACGTCCACAGTAACGACGACTACTTTGAGGTC   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    778 AAGGCTATTCAAGAACAATTGATCGTAACGTCCACAGTAACGACGACTACTTTGAGGTC   837

Qy    301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACCGGCAAGGTCAACTTTGCTGAC   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    838 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACCGGCAAGGTCAACTTTGCTGAC   897

Qy    361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCAAGAATTTTTGCTAAGCGGACAT   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT   957

Qy    421 GTGGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAAATCTGTTGATGTG   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 GTGGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAAATCTGTTGATGTG  1017

Qy    481 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACGACCAGGTCTCAAA   540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1018 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACGACCAGGTCTCAAA  1077

Qy    541 GATACTAAGCTATTGAAAACACTAGCTACGTATCGGTGACCATCACATCTCAAGAAATTACTA   600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1078 GATACTAAGCTATTGAAAACACTAGCTACGTATCGGTGACCATCACATCTCAAGAAATTACTA  1137

Qy    601 GCTCAAGCACAAAGCATTTAAACAAAAACACCACCCAGGCTATACGTTATGAACGTGAC   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1138 GCTCAAGCACAAAGCATTTAAACAAAAACACCACCCAGGCTATACGTTATGAACGTGAC  1197

Qy    661 TCCTCAATCGTCACTCACTCCATGACAATGACAATTTCCGTACGATTTACCAATCGGATCAAGAG   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1198 TCCTCAATCGTCACTCACTCCATGACAATGACAATTTCCGTACGATTTACCAATCGGATCAAGAG  1257

Qy    721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG   780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1258 TTTACTTACCGTGTTAAAAATCGGAACAAGCTTATAGAGATCAATAAAAATCTGGTCTG  1317

Qy    781 AATGAAGAAATAACAACACTGACCTGCTGAGAAATATTACGTCCTTAAAAAAAGGG   840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1318 AATGAAGAAATAAACACACTGACCTGCTGAGAAATATTACGTCCTTAAAAAAAGGG  1377

Qy    841 GAAAAGCCGTATGATCCCTTTGATGCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1378 GAAAAGCCGTATGATCCCTTTGATGCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  1437

Qy    901 GATGTCAACACCAACGAATTGCTAAAAAGCCAGCAGCTCTTAACAGCTAGCGAACGTAAC   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1438 GATGTCGATACCAACGAATTGCTAAAAAGTGCAGCAGCTCTTAACAGCTAGCGAACGTAAC  1497

Qy    961 TTTAGACTTCAGAGATTTATACGATCCGTGATAAGGCTAAACTCTTACAACAATCTC  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1498 TTTAGACTTCAGAGATTTATACGATCCGTGATAAGGCTAAACTACTCTACAACAATCTC  1557

Qy   1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGAATAATCACGATGAC  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1558 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGAATAATCACGATGAC  1617

Qy   1201 CGTTATACAGGGACACCTATACCTGATAACCCTAAGGACAAAATAA  1245
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1333 CGTTATACAGGGACACCTATACCTGATAACCCTAAGGACAAAATA  1377
```

RESULT 10
AAA37642
ID   AAA37642 standard; DNA; 1782 BP.
XX
AC   AAA37642;
XX
DT   15-SEP-2003 (revised)
DT   13-OCT-2000 (first entry)
XX
DE   Chimeric SK-FBD coding sequence.
XX
KW   Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW   plasminogen; human; fibronectin; thrombolytic therapy;
KW   cardiovascular disorder; ss.
XX
OS   Streptococcus dysgalactiae subsp. equisimilis.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   EP1024192-A2.
XX
PD   02-AUG-2000.
XX
PF   23-DEC-1999; 99EP-00310541.
XX
PR   24-DEC-1998; 98IN-DE003825.
XX
PA   (COUL) CSIR COUNCIL SCI IND RES.
XX
PI   Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI   Yadav M;
XX
DR   WPI; 2000-516032/47.
XX
PT   Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT   thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT   domains of human fibronectin.
XX
PS   Example 5; Fig 21b; 58pp; English.
XX
CC   This sequence represents a chimeric streptkinase-fibrin binding domain
CC   (SK-FBD) protein coding sequence. The invention relates to a hybrid
CC   plasminogen activator (PA) comprises a polypeptide fusion between
CC   streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC   fibrin binding regions of human fibronectin, which are from fibrin
CC   binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC   ability to bind with fibrin independently and also characteristically
CC   retains a PG activation ability which becomes evident only after a
CC   pronounced duration, or lag, after exposure of the PA to a suitable
CC   animal or human PG. The hybrid streptokinase-fibrin binding domain
CC   polypeptides are useful in thrombolytic therapy for various kinds of
CC   cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC   well as kinetics of plasminogen activation that are distinct from that of
CC   natural streptokinase in being characterised by a temporary delay, or lag
CC   of several minutes in the natural rate of the catalytic conversion of
CC   plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC   can bind tightly with fibrin in blood clots soon after introduction into
CC   the vascular system without significantly activating the circulating
CC   blood plasminogen to plasmin, thus aiding in the localisation of the
CC   plasminogen activation process to the site of pathological thrombus. This
CC   overcomes systemic plasminogen activation encountered during clinical use
CC   of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ   Sequence 1782 BP; 573 A; 407 C; 377 G; 425 T; 0 U; 0 Other;

Query Match   98.2%; Score 1222.6; DB 3; Length 1782;
Best Local Similarity 98.9%; Pred. No. 0;

```
Db   358 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACACCCAATTAGTTGTT 417
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    61 AGCGTTGCTGGTACTGTTGAGGGGGACGAATCAAGACATTAGTTCTTAAATTTTTGAAATT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   418 AGCGTTGCTGGTACTGTTGAGGGGGACGAATCAAGACATTAGTTCTTAAATTTTTGAAATC 477

Qy   121 GACTTAACATCACGACCTGCTCATGGAGGAAAGACAGAGAGGCTTAAGTCCAAAATCA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   478 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGAGGCTTAAGTCCAAAATCA 537

Qy   181 AAACCATTTGCTACTGATAGTTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   538 AAACCATTTGCTACTGATAGTTGGCGCGATGCCACATAAACTTGAGAAAGCTGACTTACTA 597

Qy   241 AAGGCTATTCAAGACACAATTGATCGTCAAGCTCCACGTAACGACGACTACTTTGAGGTC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   598 AAGGCTATTCAAGAACAATTGATCGTCAACCTCCACGTAACGACGACTACTTTGAGGTC 657

Qy   301 ATTGATTTTGCAAGGGATGCAACCATTACTGATGATCGAAACGGCAAGTCTACTTTGCTGAC 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   658 ATTGATTTTGCAAGGGATGCAACCATTACTGATGATCGAAACGGCAAGTCTACTTTGCTGAC 717

Qy   361 AAAGATGGTTCGGTAACCTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   718 AAAGATGGTTCGGTAACCTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 777

Qy   421 GTGCGCGGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   778 GTGCGCGGTTAGACCATATAAAGAAAAACCAATACAAAAACCAAGCGAAATCTGTTGATGTG 837

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACAATTTCAGACCAGGTCTCAAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   838 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACAATTTCAGACCAGGTCTCAAA 897

Qy   541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   898 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 957

Qy   601 GCTCAAGCACAAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAAGCGTGAC 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   958 GCTCAAGCACAAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAAGCGTGAC 1017

Qy   661 TCCTCAATCGTCACTCATGACAATGACAATTTCCGTACGATTTTACCAATGGATCAAGAG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1018 TCCTCAATCGTCACTCATGACAATGACAATTTCCGTACGATTTTACCAATGGATCAAGAG 1077

Qy   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1078 TTTACTTACCGTGTTAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 1137

Qy   781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1138 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1197

Qy   841 GAAAAGCCGTATGATCCCTTTGATCGCGAGTCACTTGAAACTGTTCCATCAAATACGTT 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1198 GAAAAGCCGTATGATCCCTTTGATCGCGAGTCACTTGAAACTGTTCCATCAAATACGTT 1257

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGCAGCAGGAGCAGCTCTTAACAGCTAGCGAACGTAAC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1258 GATGTCGATCACCAACGAATTGCTAAAAAGCAGCAGGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1317

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAGCTAAACTACTTCTACAACAATCTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1318 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAGCTAAACTACTTCTACAACAATCTC 1377

Qy  1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1378 GATGCTTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1437

Qy  1081 ACCAACCGTATCATAACCGTTTATAATGGCAAGCGACCCGAAGGAGAATGCTAGCTAT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1438 ACCAACCGTATCATAACCGTTTATAATGGCAAGCGACCCGAAGGAGAATGCTAGCTAT 1497
```

```
Qy  1081 ACCAACCGTATCATAACCGTTTATAATGGGCAAGCGACCGACCCGAAGGAGAGAGAATGCTAGCTAT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1618 ACCAACCGTATCATAACCGTTTATAATGGGCAAGCGACCGACCCGAAGGAGAGAGAATGCTAGCTAT 1677

Qy  1141 CATTAGCCCATGATAAAGATCGTTATACCGAAGAAGAACGAGAGAGTTTACAGCTACCTG 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1678 CATTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAGAGTTTACAGCTACCTG 1737

Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1738 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1782
```

RESULT 11
AAQ11651

ID   AAQ11651   standard; DNA; 2030 BP.
XX
AC   AAQ11651;
XX
DT   08-JUL-1991 (first entry)
XX
DE   FB-FB-SK fusion construct.
XX
KW   Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
KW   streptokinase; fusion protein; ss.
XX
OS   Staphylococcus aureus.
XX
FH   Key   Location/Qualifiers
FT   CDS   10..184
FT   /tag= a
FT   /label= FB monomer
FT   CDS   185..358
FT   /tag= b
FT   /label= FB monomer
FT   CDS   359..1601
FT   /tag= c
FT   /label= streptokinase
XX
PN   US5011686-A.
XX
PD   30-APR-1991.
XX
PF   15-NOV-1989; 89US-00437769.
XX
PR   21-SEP-1987; 87US-00099242.
XX
PA   (CREA-) CREATIVE BIOMOLEC.
XX
PI   Pang RHL;
XX
DR   WPI; 1991-140198/19.
DR   P-PSDB; AAR11829.
XX
PT   Imparting injectable fibrinolytic agent - with affinity for intravascular
PT   thrombus, by linking agent to fibrin binding domain.
XX
PS   Disclosure; Fig 5; 18pp; English.
XX
CC   The DNA encodes an FB-FB dimer linked to the streptokinase coding
CC   sequence. The FB fragment has selective affinity for fibrin, low affinity
CC   for fibrinogen, and minimal immunogenicity, imparting thrombus-targetting
CC   capability. See also AAQ11649 and AAQ11650
XX
SQ   Sequence 2030 BP; 673 A; 464 C; 406 G; 487 T; 0 U; 0 Other;

Query Match   98.2%;   Score 1222.6;   DB 2;   Length 2030;
Best Local Similarity   98.9%;   Pred. No. 0;
Matches 1231;   Conservative   0;   Mismatches   14;   Indels   0;   Gaps   0;

```
Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
         |||||||||||||||||||||||||||||||||||||||||||||||
```

Best Local Similarity 99.0%; Pred. No. 0;
Matches 1229; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy      1 ATTGCTGGACCTGAGTGGCTGCTGCTAGACGTCCATCTGTCAACACAGCCAATTAGTTGTT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATTGCTGGACCTGAGTGGCTGCTGCTAGACGTCCATCTGTCAACACAGCCAATTAGTTGTT   60

Qy     61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAAATTTTTGAAATT   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAAATTTTTGAAATC  120

Qy    121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180

Qy    181 AAACCATTTGCTACTGATAGTGGGCGATGCCACATAAACTTGAAAAAGTGACTTACTA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAACCATTTGCTACTGATAGTGGGCGATGTCACATAAACTTGAGAAAGTGACTTACTA   240

Qy    241 AAGGCTATTCAAGAACAATTGATCGTAACGTCCACAGTAACGACGAACGACTACTTTGAGGTC  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AAGGCTATTCAAGAACAATTGATCGTAACGTCCACAGTAACGACGAACGACTACTTTGAGGTC  300

Qy    301 ATTGATTTTGCAAGCGATGCAACCATTACTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATTGATTTTGCAAGCGATGCAACCATTACTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC  360

Qy    361 AAAGATGGTTCGGTAACCTTGCCGACCCACCTGTCCCAAGAATTTTGCTAAGCCGGACAT   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAGATGGTTCGGTAACCTTGCCGACCCACCTGTCCCAAGAATTTTGCTAAGCCGGACAT   420

Qy    421 GTGCGCGTTAGACCATATAAAGAGAAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGGCGCGTTAGACCATATAAAGAGAAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG  480

Qy    481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA   540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA   540

Qy    541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACCACCATCACATCTCAAGAATTACTA   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACCACCATCACATCTCAAGAATTACTA   600

Qy    601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC   660

Qy    661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG   720

Qy    721 TTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG   780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TTTACTTACCGTGTTAAAATCGGGAACAAGCTTATAGGATCAATAAAAATCTGGTCTG   780

Qy    781 AATGAAGAAATAAACAACACTGACTGATCCTCGAGAAATATTACGTCCTTAAAAAAGGG  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AATGAAGAAATAAACAACACTGACTGATCCTCGAGAAATATTACGTCCTTAAAAAAGGG  840

Qy    841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTTGTTCACCATCAAAATACGTT   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTTGTTCACCATCAAAATACGTT   900

Qy    901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC   960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCAGCCGAACGTAAC   960

Qy    961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC  1020

Qy   1021 GATGCTTTTGGTATTATGATGGACTATACCTTAACTGAACAGTAGAGGAATAATCACGATGAC  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1498 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG  1557

Qy   1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA  1245
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1558 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA  1602
```

RESULT 12

AAX80492
ID   AAX80492 standard; cDNA; 1242 BP.
XX
AC   AAX80492;
XX
DT   17-OCT-2003 (revised)
DT   26-AUG-1999 (first entry)
XX
DE   Streptococcus equisimilis native streptokinase encoding cDNA.
XX
KW   Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW   nSK; rSK; bacterial; blood clot; thrombotic condition;
KW   myocardial infarction; venous thrombosis; pulmonary embolism;
KW   cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
XX
OS   Streptococcus dysgalactiae subsp. equisimilis.
XX
PN   WO9931247-A1.
XX
PD   24-JUN-1999.
XX
PF   15-DEC-1998; 98WO-US026694.
XX
PR   15-DEC-1997; 97US-0069497P.
XX
PA   (HARD) HARVARD COLLEGE.
XX
PI   Reed GL;
XX
DR   WPI; 1999-395183/33.
DR   P-PSDB; AAY24794.
XX
PT   N-terminally deleted streptokinase.
XX
PS   Claim 44; Page 58-60; 73pp; English.
XX
CC   The present invention describes an isolated bacterial protein that
CC   induces fibrin-dependent plasminogen activation in a pharmaceutical
CC   composition for dissolving blood clots. Also described are: (1) a
CC   composition comprising an isolated modified streptokinase, the
CC   modification being removal of amino acid residues in the amino terminus;
CC   (2) a method for dissolving a blood clot in a subject, comprising
CC   administering to the subject a fibrin-dependent streptokinase protein; a
CC   nucleic acid (I) encoding a modified bacterial streptokinase; (3) an
CC   expression vector comprising (I); and (4) a host cell transformed with
CC   the expression vector of (3). The pharmaceutical composition comprising a
CC   bacterial fibrin-dependent plasminogen activator is useful for dissolving
CC   blood clots in patients with a thrombotic condition, e.g. myocardial
CC   infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
CC   graft thrombosis and arterial thrombosis. The modified streptokinase can
CC   also be used in non-human mammals. Streptokinase activation of
CC   plasminogen is at least 10-fold, preferably 100-fold greater in the
CC   presence of fibrin than in the absence of fibrin. The modified
CC   streptokinase has at least one amino acid substitution that inactivates a
CC   substrate site for proteolytic cleavage. This reduces the rate of
CC   degradation of the streptokinase at least two-fold. The present sequence
CC   encodes native streptokinase (nSK). (Updated on 17-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 1242 BP; 424 A; 267 C; 237 G; 314 T; 0 U; 0 Other;

Query Match   98.1%; Score 1221.2; DB 2; Length 1242;

```
Db           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       7     ATTGCTGGACCTGAGTGGCTGCTAGACCGTTCATCTGTCAACAACAGCCAATTAGTTGTT    66

Qy     61    AGCGGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT   120
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     67    AGCGGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC   126

Qy     121   GACCTAACATCACGACCTGCTCATGAGGAGGAAGAACAGAGCAGGCAAGGCTTAAGTCCAAAATCA   180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     127   GATCTAACATCACGACCTGCTCATGAGGAGGAAGAACAGAGCAGGCAAGGCTTAAGTCCAAAATCA   186

Qy     181   AAACCATTTGCTACTGATAGTTGGCGGATGCCACATAAACTTGAAAAAGCTGACTTACTA   240
             ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     187   AAACCATTTGCTACTGATAGTTGGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA   246

Qy     241   AAGGCTATTCAAGAACAATTGATCGTCAACGTCCACAGTAACGACGACTACTTTGAGGTC   300
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     247   AAGGCTATTCAAGAACAATTGATCGTCAACGTCCACAGTAACGACGACTACTTTGAGGTC   306

Qy     301   ATTGATTTTGCAAGCGATGCAACCATTACTGAAAAACGGCAAGGTCTACTTTGCTGAC   360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     307   ATTGATTTTGCAAGCGATGCAACCATTACTGAAAAACGGCAAGGTCTACTTTGCTGAC   366

Qy     361   AAAGATGGTTCGGTAACCTTGCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT   420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     367   AAAGATGGTTCGGTAACCTTGCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT   426

Qy     421   GTGCGCGGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG   480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     427   GTGCGCGGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG   486

Qy     481   GAATATACTGTACAGTTACTCCCTTAAACCTGATGCAGATTTCAGCACCAGTCTCAAA   540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     487   GAATATACTGTACAGTTACTCCCTTAAACCCTGATGCAGATTTCAGCACCAGTCTCAA   546

Qy     541   GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA   600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     547   GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA   606

Qy     601   GCTCAAGCACACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC   660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     607   GCTCAAGCACACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC   666

Qy     661   TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG   720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     667   TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG   726

Qy     721   TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG   780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     727   TTTACTTACCGTGTTAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAAATCTGGTCTG   786

Qy     781   AATGAGAAATAAACACACTGACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG   840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     787   AATGAGAAATAAACACACTGACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG   846

Qy     841   GAAAAGCCGTATGATCCCTTTGATCGCGAGTCACTTGAAACTGTTCACCATCAAATACGTT   900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     847   GAAAAGCCGTATGATCCCTTTGATCGCGAGTCACTTGAAACTGTTCACCATCAAATACGTT   906

Qy     901   GATGTCAACACCAACGAATTGCTAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC   960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     907   GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC   966

Qy     961   TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACACACAATCTC   1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     967   TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACACACAATCTC   1026

Qy     1021  GATGCTTTTGGTATTATGGACTATACCTTAACTGAGAAAAGTAGAGGAATAATCACGATGAC   1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1027  GATGCTTTTGGTATTATGGACTATACCTTAACTGAGAAAAGTAGAGGATAATAATCACGATGAC   1086

Qy     1081  ACCAACCGTATCATAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAGGAATGCTAGCTAT   1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     1021  GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC   1080

Qy     1081  ACCAACCGTATCATAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAGGAATGCTAGCTAT   1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081  ACCAACCGTATCATAACCGTTTATATGGCAAGCGACCCGAAGGAGAGAGGATGCTAGCTAT   1140

Qy     1141  CATTTAGCTGCTATGAAAGATCGTTATAAGGATCGTTATACCCGGAAGAAGTTTACAGCTACCTG   1200
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141  CATTTAGCCTGATAAAGATCGTTATACCCGGAAGAAGTTTACAGCTACCTG   1200

Qy     1201  CGTTATACAGGGACCACCTATACCTGATAACCTAACGACAAA   1242
             ||||||||||||||||||||||||||||||||||||||||||
Db     1201  CGTTATACAGGGACCACCTATACCTGATAACCCTAACGACAAA   1242
```

RESULT 13
ABA05546

ID   ABA05546 standard; cDNA; 1254 BP.
XX
AC   ABA05546;
XX
DT   26-FEB-2002 (first entry)
XX
DE   Streptokinase cDNA.
XX
KW   Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;
KW   vasodilator; thrombolytic; angina; myocardial infarction; stroke;
KW   gene therapy; maxadilan; ss.
XX
OS   Unidentified.
XX
PN   WO200185100-A2.
XX
PD   15-NOV-2001.
XX
PF   10-MAY-2001; 2001WO-US015209.
XX
PR   11-MAY-2000; 2000US-00569920.
XX
PA   (GEHO) GEN HOSPITAL CORP.
XX
PI   Reddy VB, Lerner E;
XX
DR   WPI; 2002-062184/08.
XX
PT   New fusion protein or conjugate, useful for treating unstable angina,
PT   acute myocardial infarction or stroke, comprises a vasodilator
PT   polypeptide and a thrombolytic polypeptide, or active fragments of the
PT   polypeptides.
XX
PS   Example 1; Fig 2; 37pp; English.
XX
CC   The invention relates to a fusion protein or a conjugate comprising a
CC   vasodilator polypeptide, or its active fragment, and a thrombolytic
CC   polypeptide or its active fragment. The protein is useful for treating a
CC   subject suffering from a partially or totally occluded blood vessel,
CC   causing unstable angina, acute myocardial infarction or stroke. The
CC   polynucleotide encoding the polypeptide is useful in gene therapy. The
CC   vasodilatory action of the protein allows for the use of lower doses of a
CC   thrombolytic while maintaining the clot dissolving effectiveness of the
CC   thrombolytic, and the use of lower doses of the thrombolytic reduces
CC   associated side effects. The present sequence is the streptokinase cDNA
CC   used in the construction of a Maxadilan-Streptokinase fusion protein.
CC   Maxadilan is a vasodilator peptide produced by the salivary gland of the
CC   New World sand fly
XX
SQ   Sequence 1254 BP; 426 A; 271 C; 241 G; 316 T; 0 U; 0 Other;

Query Match   98.1%; Score 1221.2; DB 6; Length 1254;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1229; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy   1   ATTGCTGGACCTGAGTGGCTGCTAGACCGTTCATCTGTCAACAACAGCCAATTAGTTGTT   60

```
Db     1087  ACCAACCGTATCATACACCGTTTATATGGGCAAGCGCAAGGAGAGGAGAATGCTAGCTAT 1146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1141  CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1147  CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1206

Qy     1201  CGTTATACAGGGACACCTATACCTGATAACCTAACGACAAA 1242
             |||||||||||||||||||||||||||||||||||||||||
Db     1207  CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA 1248

RESULT 14
AAX80497
ID   AAX80497 standard; cDNA; 2385 BP.
XX
AC   AAX80497;
XX
DT   26-AUG-1999  (first entry)
XX
DE   Streptokinase and maltose binding protein fusion protein encoding cDNA.
XX
KW   Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW   nSK; rSK; bacterial; blood clot; thrombotic condition;
KW   myocardial infarction; venous thrombosis; pulmonary embolism;
KW   cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
XX
OS   Streptococcus dysgalactiae subsp. equisimilis.
OS   Synthetic.
XX
PN   WO9931247-A1.
XX
PD   24-JUN-1999.
XX
PF   15-DEC-1998;  98WO-US026694.
PR   15-DEC-1997;  97US-0069497P.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Reed GL;
XX
DR   WPI; 1999-395183/33.
DR   P-PSDB; AAY24797.
XX
PT   N-terminally deleted streptokinase.
XX
PS   Example; Page 45-48; 73pp; English.
XX
CC   The present invention describes an isolated bacterial protein that
CC   induces fibrin-dependent plasminogen activation in a pharmaceutical
CC   composition for dissolving blood clots. Also described are: (1) a
CC   composition comprising an isolated modified streptokinase, the
CC   modification being removal of amino acid residues in the amino terminus;
CC   (2) a method for dissolving a blood clot in a subject, comprising
CC   administering to the subject a fibrin-dependent streptokinase protein; a
CC   nucleic acid (I) encoding a modified bacterial streptokinase; (3) an
CC   expression vector comprising (I); and (4) a host cell transformed with
CC   the expression vector of (3).  The pharmaceutical composition comprising a
CC   bacterial fibrin-dependent plasminogen activator is useful for dissolving
CC   blood clots in patients with a thrombotic condition, e.g. myocardial
CC   infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
CC   graft thrombosis and arterial thrombosis. The modified streptokinase can
CC   also be used in non-human mammals. Streptokinase activation of
CC   plasminogen is at least 10-fold, preferably 100-fold greater in the
CC   presence of fibrin than in the absence of fibrin. The modified
CC   streptokinase has at least one amino acid substitution that inactivates a
CC   substrate site for proteolytic cleavage. This reduces the rate of
CC   degradation of the streptokinase at least two-fold. The present sequence
CC   encodes a streptokinase and maltose binding protein fusion protein from
CC   an example of the present invention
XX
SQ   Sequence 2385 BP; 745 A; 547 C; 549 G; 544 T; 0 U; 0 Other;
```

```
Query Match             98.1%;  Score 1221.2;  DB 2;  Length 2385;
Best Local Similarity   99.0%;  Pred. No. 0;
Matches 1229;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      1  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCCATCTGTCAACAACAGCCAATTAGTTGTT 1203

Qy     61  AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTTGAAATT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1204  AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTTGAAATC 1263

Qy    121  GACTCAACAATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1264  GATCTAACAATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 1323

Qy    181  AAACCATTTGCTACTGATAGTGGGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324  AAACCATTTGCTACTGATAGTGGGCGATGTCACATAAACTTGAAAAAGCTGACTTACTA 1383

Qy    241  AAGGCTATTCAAGAGACAATTGATGCTAACGTCCCACAGTAACGACGACTACTTTGAGGTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384  AAGGCTATTCAAGAGACAATTGATCGTAACGTCCCACAGTAACGACGACTACTTTGAGGTC 1443

Qy    301  ATTGATTTTGCAAGCGATGCAACCATTACTACTGATCGATCGAAACGCAAGGTCTACTTGCTGAC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444  ATTGATTTTGCAAGCGATGCAACCATTACTACTGATCGATCGAAACGCAAGGTCTACTTTGCTGAC 1503

Qy    361  AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504  AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT 1563

Qy    421  GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1564  GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 1623

Qy    481  GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCACAGGTCTCAAA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1624  GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCACAGGTCTCAAA 1683

Qy    541  GATACTAAGCTATTGAAAACACTAGCTATCGTGACCACCATCACATCTCAAGAATTACTA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1684  GATACTAAGCTATTGAAAACACTAGCTATCGGTGACCACCATCACATCTCAAGAATTACTA 1743

Qy    601  GCTCAAGCACAAAGCATTTTAAACAAAACCACCCAGGCTATACGATTTATGAAACGTGAC 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1744  GCTCAAGCACAAAGCATTTTAAACAAAAACCACCCAGGCTATACGATTTATGAAACGTGAC 1803

Qy    661  TCCTCAATCGTCACTCATGACAATGACAATTTCCGTACGATTTTACCAATGATCAAGAG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1804  TCCTCAATCGTCACTCATGACAATGACAATTTCCGTACGATTTTACCAATGATCAAGAG 1863

Qy    721  TTTACTTACCATGTCAAAAATCGGAAAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1864  TTTACTTACCGTGTTAAAAATCGGAAAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 1923

Qy    781  AATGAAGAAATAAACAACACTGACCTGACTCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1924  AATGAAGAAATAAACAACACTGACCTGACTCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1983

Qy    841  GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1984  GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 2043

Qy    901  GATGTCAACACCACGAATTGCTAAAAAGCGGAGCAGCTCTTAAACAGTCTAGCGAACGTAAC 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2044  GATGTCGATACCACGAATTGCTAAAAAGTGAGCAGCTCTTAAACAGTCTAGCGAACGTAAC 2103

Qy    961  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2104  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 2163

Qy   1021  GATGCGTTTGGTATTATGGACTATACCTTAACTGCAGAAAGTAGAGGATAATCACGATGAC 1080
```

```
Best Local Similarity 99.0%;   Pred. No. 0;                Gaps       0;
Matches 1229;   Conservative      0;   Mismatches   13;   Indels     0;

Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTTCATCGTCAACAAGCCAATTAGTTGTT       60
         |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  5935 ATTGCTGGACCTGAGTGGCTGCTAGACCGTTCATCGTCAACAACAGCCAATTAGTTGTT   5994

Qy    61 AGCGTTGCTGTCGTTGAGGGGACGAATCAAGACATTAGTCTAAATTTTTGAAATT      120
         ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  5995 AGCGTTGCTGTCGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC     6054

Qy   121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6055 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 6114

Qy   181 AAACCATTTGCTACTGATAGTGGCGCGGATGCGCCACTAAAACTTGAAAAAGCTGACTTACTA  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6115 AAACCATTTGCTACTGATAGTGGCGCGGATGTGCCACTAAAACTTGAGAAAGCTGACTTACTA 6174

Qy   241 AAGGCTATTCAAGAACAATTGATCGTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6175 AAGGCTATTCAAGAACAATTGATCGTAACGTCCACAGTAACGACGACTACTTTGAGGTC 6234

Qy   301 ATTGATTTGCAAGCGATGCAACCATTACTAGTCGAAACGGCAAGGTCTACTTGCTGAC  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6235 ATTGATTTTGCAAGCGATGCAACCATTACTAGTCGAAACGGCAAGGTCTACTTGCTGAC 6294

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6295 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGGACAT 6354

Qy   421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6355 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG 6414

Qy   481 GAATATACTGTACAGTTTACTCCCTAAACCCTGATGACGATTTCAGACCAGTCTCAAA  540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6415 GAATATACTGTACAGTTTACTCCCTAAACCCTGATGACGATTTCAGACCAGTCTCAAA 6474

Qy   541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6475 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 6534

Qy   601 GCTCAAGCACAAAGCATTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6535 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC 6594

Qy   661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6595 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG 6654

Qy   721 TTTACTTACCATGTCAAAAATCGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6655 TTTACTTACCGTGTTAAAAATCGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 6714

Qy   781 AATGAAGAAATAAACAACACTGACCTTGCTGAGAAATATTACGTCCTTAAAAAAGGG  840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6715 AATGAAGAAATAAACAACACTGACCTTCTGAGAAATATTACGTCCTTAAAAAAGGG 6774

Qy   841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6775 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 6834

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGTCTAGCGAACGTAAC  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6835 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 6894

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTTCTACAACAATCTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6895 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTTCTACAACAATCTC 6954

Qy  1021 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAAAGTAGAGGATAATCACCGATGAC 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2164 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAAAGTAGAGGATAATCACCGATGAC 2223

Qy  1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCCACCCGAAGGAGGAGAATGCTAGCTAT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2224 ACCAACCGTATCATAACCGTTTATATGGGCAAGCCACCCGAAGGAGGAGAATGCTAGCTAT 2283

Qy  1141 CATTTAGCCTATGATAAAGATCGTTATACCGTATACGTATACGAGAGTTTACAGTCACCTG 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2284 CATTTAGCCTATGATAAAGATCGTTATACCGTATACGTATACGAGAGTTTACAGCTACCTG 2343

Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA 1242
         ||||||||||||||||||||||||||||||||||||||||||
Db  2344 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA 2385
```

Query Match   98.1%; Score 1221.2; DB 6; Length 8893;

```
Db        6955 GATGCTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 7014

Qy        1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGGAATGCTAGCTAT 1140
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7015 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGGAATGCTAGCTAT 7074

Qy        1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7075 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 7134

Qy        1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA 1242
               ||||||||||||||||||||||||||||||||||||||||||
Db        7135 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA 7176
```

Search completed: May 6, 2004, 03:54:39
Job time : 555 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: May 6, 2004, 03:22:22 ; Search time 5089 Seconds
(without alignments)
10603.670 Million cell updates/sec

Title: US-09-882-509-3
Perfect score: 1245
Sequence: 1 attgctgacctgagtggct.........ataaccctaacgacaaataa 1245

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:_
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

SUMMARIES

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1245	100.0	1245	1	S46536	S46536 SKC-2=strep
2	1241.8	99.7	1257	6	A20015	A20015 SEQ ID NO:
3	1241.8	99.7	1257	6	I13203	I13203 Sequence 25
4	1241.8	99.7	1317	6	A20009	A20009 SEQ ID NO:
5	1241.8	99.7	1317	6	I13197	I13197 Sequence 18
6	1241.8	99.7	1335	6	A20006	A20006 SEQ ID NO:
7	1241.8	99.7	1335	6	I13194	I13194 Sequence 14
8	1241.8	99.7	1458	6	A20027	A20027 SEQ ID NO:
9	1241.8	99.7	1458	6	I13215	I13215 Sequence 42
10	1241.8	99.7	1512	6	A20016	A20016 SEQ ID NO:
11	1241.8	99.7	1512	6	I13204	I13204 Sequence 27
12	1241.8	99.7	2589	6	I13209	I13209 Sequence 34
13	1238.8	99.5	2589	6	A20030	A20030 SEQ ID NO:
14	1238.8	99.5	1467	6	I13218	I13218 Sequence 46
15	1238.8	99.5	2588	6	A20021	A20021 SEQ ID NO:
16	1225.8	98.5	1311	1	AY368335	AY368335 Streptoco
17	1224.2	98.3	1473	1	SGSKG	X13400 Streptococc
18	1224.2	98.3	2568	1	STRSKC	K02986 Streptococc
19	1224.2	98.3	2568	1	A04926	A04926 S.equisimil
20	1224.2	98.3	8931	1	SEDEXB	X72832 S.equisimil
21	1221.2	98.1	1242	6	AR144000	AR144000 Sequence
22	1221.2	98.1	2385	6	AR143998	AR143998 Sequence
23	1219.4	97.9	2568	6	E00522	E00522 DNA fragmen
24	1217.8	97.8	1401	6	I05204	I05204 Sequence 5
25	1210.2	97.2	2566	6	AR068768	AR068768 Sequence
26	1206.4	96.9	1209	6	AR175891	AR175891 Sequence
27	1206.4	96.9	1209	6	AX030315	AX030315 Sequence
28	1203.8	96.7	1245	6	AR175892	AR175892 Sequence
29	1203.8	96.7	1245	6	AX030316	AX030316 Sequence
30	1192.2	95.8	1407	6	E01413	E01413 DNA sequenc
31	1116	89.6	1122	6	AR175893	AR175893 Sequence
32	1116	89.6	1122	6	AX030317	AX030317 Sequence
33	1116	89.6	1158	6	AR175894	AR175894 Sequence
34	1116	89.6	1158	6	AX030318	AX030318 Sequence
35	1104.2	88.7	1323	1	AY234130	AY234130 Streptoco
36	1104.2	88.7	50354	1	AE014169	AE014169 Streptoco
37	1104.2	88.7	323825	1	AP005146	AP005146 Streptoco
38	1103.4	88.6	1119	6	I13206	I13206 Sequence 30
39	1103.2	88.6	2252	6	A20025	A20025 SEQ ID NO:
40	1103.2	88.6	2253	6	I13213	I13213 Sequence 39
41	1102.6	88.6	1323	1	AY234129	AY234129 Streptoco
42	1102.6	88.6	1323	1	AY234137	AY234137 Streptoco
43	1096.2	88.0	1323	1	AY234128	AY234128 Streptoco
44	1096.2	88.0	10586	1	AE006620	AE006620 Streptoco
45	1094.6	87.9	1473	1	SPSKA	X13399 Streptococc

ALIGNMENTS

RESULT 1
S46536
LOCUS   S46536   1245 bp   DNA   linear   BCT 08-MAY-1993
DEFINITION   SKC-2=streptokinase [Streptococcus equisimilis, group C, ATCC 9542,
Genomic, 1245 nt].
ACCESSION   S46536
VERSION   S46536.1   GI:257196
KEYWORDS   .
SOURCE   Streptococcus dysgalactiae subsp. equisimilis
ORGANISM   Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE   1 (bases 1 to 1245)
AUTHORS   Estrada,M.P., Hernandez,L., Perez,A., Rodriguez,P., Serrano,R.,
Rubiera,R., Pedraza,A., Padron,G., Antuch,W., de la Fuente,J.

Pred. No. is the number of results predicted by chance to have a

```
                    et.al.
TITLE               High level expression of streptokinase in Escherichia coli
JOURNAL             Biotechnology (N.Y.) 10 (10), 1138-1142 (1992)
MEDLINE             93000998
PUBMED              1368792
REMARK              GenBank staff at the National Library of Medicine created this
                    entry [NCBI gibbsq 115306] from the original journal article.
                    This sequence comes from Fig. 2A.
FEATURES            Location/Qualifiers
     source         1..1245
                    /organism="Streptococcus dysgalactiae subsp. equisimilis"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:119602"
                    /note="group: C"
     gene           1..1245
                    /gene="SKC-2"
     CDS            1..1245
                    /gene="SKC-2"
                    /note="streptokinase; This sequence comes from Fig. 2B;
                    Mature protein lacks initial Met"
                    /codon_start=1
                    /product="SKC-2"
                    /protein_id="AAC60418.1"
                    /db_xref="GI:257197"
                    /translation="IAGPEWLLDRPSVNNSQLVVSVAGTVEGTNQDISLKFFEIDLTS
                    RPAHGGKTEQGLSPKSKPFATDSGAMPHKLEKADLLKAIQEQLIANVHSNDDYFEVID
                    FASDATITDRNGKYFADKDGSVYFLPTQPVQEFLLSGHVRVRPYKEKPIQNQAKSVDV
                    EYTVQFTPLNPDDDFRPGLKDTKLLKTLAIGDTITSQRLLAQASILNKTHPGYTIYE
                    RDSSIVTHDNDIFRTILPMDQEFTYHVKNREQAYEINKKSGLNEEINNTDLISEKYYV
                    LKKGEKPYDPFDRSHLKLFTIKYVDVNTNELLKSEQLITASERNLDFRDLYDPRDKAK
                    LLYNNLDAPGIMDYTLTGKVEDNHDDTNRIITVYMGKRPEGENASYHLAYDKDRYTEE
                    EREVYSYLRYTGTPIPDNPNDK"
ORIGIN

Query Match         100.0%;   Score 1245;  DB 1;  Length 1245;
Best Local Similarity 100.0%;  Pred. No. 5.2e-297;
Matches 1245;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   60

Qy       61 AGCGGTTGCTGTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 AGCGGTTGCTGTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATT  120

Qy      121 GACCTAACATCACGACACCTGCTCATGGAGGAAAGACAAGAGCAAGGCTTAAGTCCAAAATCA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GACCTAACATCACGACACCTGCTCATGGAGGAAAGACAAGAGCAAGGCTTAAGTCCAAAATCA  180

Qy      181 AAACCATTTGCTACTGATAGTGGCGCGATGGCCCACATAAACTTGAAAAAGCTGACTTACTA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 AAACCATTTGCTACTGATAGTGGCGCGATGGCCCACATAAACTTGAAAAAGCTGACTTACTA  240

Qy      241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300

Qy      301 ATTGATTTTGCAAGCGATGCAACATTACTGATCGAATGCAACATTACTGATCGAATGCTGAC  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ATTGATTTTGCAAGCGATGCAACATTACTGATCGAATGCAACATTACTGATCGAATGCTGAC  360

Qy      361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420

Qy      421 GTGCGCGGTTAGACATATAAAGAAAAACAATACAAATCAAGCGAAATCGTTGATGTG  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GTGCGCGGTTAGACATATAAAGAAAAACAATACAAATCAAGCGAAATCGTTGATGTG  480

Qy      481 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACGATTTCAGACCAGGTCTCAA  540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACGATTTCAGACCAGGTCTCAA  540
```

```
Qy      541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA  600

Qy      601 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC  660

Qy      661 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTACCAATGGATCAAGAG  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTACCAATGGATCAAGAG  720

Qy      721 TTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 TTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG  780

Qy      781 AATGAGAAATAAACACACTGACCTGATCCTCGAGAAATATTACGTCCTTAAAAAGGG  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 AATGAGAAATAAACACACTGACCTGATCCTCGAGAAATATTACGTCCTTAAAAAGGG  840

Qy      841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  900

Qy      901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTTCTTAACAGCTAGCGAACGTAAC  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTTCTTAACAGCTAGCGAACGTAAC  960

Qy      961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTTCTACACAATCTC  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTTCTACACAATCTC  1020

Qy     1021 GATGCTTTTGGTTATTATGGACTATACCTTAACTGGCAGAATCACGATGAC  1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 GATGCTTTTGGTTATTATGGACTATACCTTAACTGGCAGAATCACGATGAC  1080

Qy     1081 ACCAACCGTATCATACACCGTTATATGGGCAGCAAGCACCCGAAGGAGGAGAAGAATGCTAGCTAT  1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 ACCAACCGTATCATACACCGTTATATGGGCAGCAAGCACCCGAAGGAGGAGAAGAATGCTAGCTAT  1140

Qy     1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAAGTTTACAGCTACCTG  1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAAGTTTACAGCTACCTG  1200

Qy     1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA  1245
            |||||||||||||||||||||||||||||||||||||||||||||
Db     1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA  1245
```

```
RESULT 2
A20015
LOCUS               A20015                1257 bp    DNA        linear   PAT 14-JUL-1995
DEFINITION          SEQ ID NO: 23; Nucleotide sequence for methionol-streptokinase
                    fusion protein.
ACCESSION           A20015
VERSION             A20015.1   GI:1247848
KEYWORDS            .
SOURCE              synthetic construct
  ORGANISM          synthetic construct
                    artificial sequences.
REFERENCE           1 (bases 1 to 1257)
  AUTHORS           .
  TITLE             PROTEINS AND NUCLEIC ACIDS
  JOURNAL           Patent: WO 9109125-A 23 27-JUN-1991;
FEATURES            Location/Qualifiers
     source         1..1257
                    /organism="synthetic construct"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32630"
     CDS            4..1251
                    /codon_start=1
                    /transl_table=11
```

/product="methionyl-streptokinase fusion protein"
/protein_id="CAA01486.1"
/db_xref="GI:1247849"
/db_xref="REMTREMBL:CAA01486"
/translation="MIAGPEWILDRPSVNNSQLVVSVAGTVEGTNQDISLKFFEIDLT
SRPAHGGKTEQGLSPKSKPFATDSGAMPHKLEKADLLKAIQEQLIANVHSNDDYFEVI
DFASDATTTDRNGKVYFADKDGSVTLPTQPVQBFLLSGHVRVRPYKEKPIQNQAKSVD
VEYTVQEFTPLNPDDDFRPGLKDTKLLKTLAIGDTITSQELLAQOAQGSILNKTHPGYTIY
ERDSSIVTHDNDIFRTILPMDQEFTYHVKNRRQAYEINKKSGLNEEINNTDLISEKY
VLKKGEKPYDPFDRSHLFKLFTIKYVDVTNELLKSEQLLTASRNLDFRDLYDPRDKA
KLLYNNLDAFGIMDYTLTGKVEDNHDDTNRIITVYMGKRPEGENASYHLAYDKDRYTE
EREVYSYLRYTGTPIPDNPNDK"

ORIGIN

Query Match   99.7%;   Score 1241.8; DB 6; Length 1257;
Best Local Similarity 99.8%; Pred. No. 3.2e-296;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 ATTGCTGGACCTGAGTGGCTGCTAGACGTTCCATCTGTCAACACAGCCAATTAGTTGTT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7 ATTGCTGGACCTGAGTGGCTGCTAGACGTTCCATCTGTCAACACAGCCAATTAGTTGTT 66

Qy   61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATT 126

Qy   121 GACCTAACATCACGACCTGCTCATGGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 GACCTAACATCACGACCTGCTCATGGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 186

Qy   181 AAACCATTTGCTACTGATAGTGGCGCGATGGCGCCAATAAACTTGAAAAAGTCGACTTACTA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 AAACCATTTGCTACTGATAGTGGCGCGATGGCGCCAATAAACTTGAAAAAGTCGACTTACTA 246

Qy   241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 306

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 366

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGGACAT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGGACAT 426

Qy   421 GTGCGCGGTTAGACAAACCAATAAAAGAAAAAACCAATACAAAATCAAGCGAAATCGTTGATGTG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 GTGCGCGGTTAGACAAACCAATAAAAGAAAAAACCAATACAAAATCAAGCGAAATCGTTGATGTG 486

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACGACCAGGTCTCAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACGACCAGGTCTCAA 546

Qy   541 GATACTAAGGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 GATACTAAGGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 606

Qy   601 GCTCAAGCACAAAAGCAATTTAAACAAAACCCAGGCTATACGTATGATTTATGAACGTGAC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607 GCTCAAGCACAAAAGCAATTTAAACAAAACCCAGGCTATACGTATGATTTATGAACGTGAC 666

Qy   661 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 726

Qy   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCATTTTCGTACGATCAATAAAAATCGGTTCTG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727 TTTACTTACCATGTCAAAAATCGGGAACAAGCATTTTCGTACGATCAATAAAAATCGGTTCTG 786

Qy   781 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCCCTTAAAAAAGGG 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   787 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCCCTTAAAAAAGGG 846

Qy   841 GAAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAAACTGTTCACCATCAAAATACGTT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   847 GAAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAAACTGTTCACCATCAAAATACGTT 906

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGGGAGCAGCTCTTAACAGCTAGGCAACGTAAC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   907 GATGTCAACACCAACGAATTGCTAAAAAGGGAGCAGCTCTTAACAGCTAGGCAACGTAAC 966

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAACTACTCTCACACAATCTC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   967 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAACTACTCTCACACAATCTC 1026

Qy   1021 GATGCTTTTGGTATTATGGGACTATACCTTAACTGAAAAAGTAGAGATAATCACGATGAC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1027 GATGCTTTTGGTATTATGGGACTATACCTTAACTGAAAAAGTAGAGATAATCACGATGAC 1086

Qy   1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAAGAATGCTAGCTAT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1087 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAAGAATGCTAGCTAT 1146

Qy   1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1147 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1206

Qy   1201 CGTTATACAGGGACACCCTATACCTGATAACCCTGATAACCCTAACGACAAATAA 1245
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1207 CGTTATACAGGGACACCCTATACCTGATAACCCTGATAACCCTAACGACAAATAA 1251

RESULT 3
I13203
LOCUS   I13203   1257 bp   DNA   linear   PAT 26-JUL-1995
DEFINITION Sequence 25 from patent US 5434073.
ACCESSION   I13203
VERSION   I13203.1 GI:910551
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1257)
  AUTHORS   Dawson,K., Hunter,M.G. and Czaplewski,L.G.
  TITLE   Fibrinolytic and anti-thrombotic cleavable dimers
  JOURNAL   Patent: US 5434073-A 25 18-JUL-1995;
FEATURES   Location/Qualifiers
     source   1..1257
                     /organism="unknown"
                     /mol_type="unassigned DNA"

ORIGIN

Query Match   99.7%;   Score 1241.8; DB 6; Length 1257;
Best Local Similarity 99.8%; Pred. No. 3.2e-296;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTAGTTGTT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTAGTTGTT 66

Qy   61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTCAAGACATTGAACGTGAC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTCAAGACATTGAACGTGAC 126

Qy   121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGAAAGACAGAGCAAGAG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGAAAGACAGAGCAAGAG 186

Qy   181 AAACCATTTGCTACTGATAGTGCGCGATGGTGCGCGATGGTGCGCCACATAAACTTGGTCTG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 AAACCATTTGCTACTGATAGTGCGCGATGGTGCGCGATGGTGCGCCACATAAACTTGGTCTG 246

Qy   241 AAGGCTATTCAAGAACAATCAAGAACAATCGGGATGCCACATAAACTTGACTTACTA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 AAGGCTATTCAAGAACAATCAAGAACAATCGGGATGCCACATAAACTTGACTTACTA 306

SOURCE   synthetic construct
ORGANISM   synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1317)
AUTHORS   .
TITLE   PROTEINS AND NUCLEIC ACIDS
JOURNAL   Patent: WO 9109125-A 17 27-JUN-1991;
FEATURES   Location/Qualifiers
Source   1..1317
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
CDS   4..1311
            /note="unnamed protein product; ompAL fused to
            streptokinase gene"
            /codon_start=1
            /transl_table=11
            /protein_id="CAA01485.1"
            /db_xref="GI:1247842"
            /db_xref="REMTREMBL:CAA01485"
            /translation="MKKTAIAIAVALAGFATVAQAIAGPEWLLDRPSVNNSQLVVSVA
            GTVEGTNQDISLKFFEIDLTSRPAHGGKTEQGLSPKSKPFATDGSGAMPHKLEKADLLK
            AIQEQLIANVHSNDDYPEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEFLLSG
            HVRVRPYKEKPIQNQAKSVDVEYTVQFTPLNPDDDFRPGLKDTKLLKTLAIGDTITSQ
            ELLAQAQSILNKTHPGYTIYERDSSIVTHDNDIFRTILPMDQEFTYHVKNREQAYEIN
            KKSGLNEEINNTDLISEKYYVLKKGEKPYDPFDRSHLKLFTIKYVDVNTNELLKSSEQL
            TASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVDNHDDTNRIITVMGK
            RPEGRNASYHLAYDKDRYTEBEREVYSYLRYTGTPIPDNPNDK"

ORIGIN

Query Match   99.7%;   Score 1241.8; DB 6;   Length 1317;
Best Local Similarity   99.8%; Pred. No. 3.2e-296;
Matches 1243;   Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT   60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT   126

Qy   61 AGCGGTGCTGGTACTGTTGAGGGGACAGAATCAAGACATTAGTCTTAAATTTTTTGAAATT   120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 AGCGGTGCTGGTACTGTTGAGGGGACAGAATCAAGACATTAGTCTTAAATTTTTTGAAATT   186

Qy   121 GACCTAACATCACGACCTGCTCATGGAGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA   180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 GACCTAACATCACGACCTGCTCATGGAGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA   246

Qy   181 AAACCATTGCTACTGAGTAGTGGCGCGATGGCCACATAAAACTTGAAAAAGCTGACTTACTA   240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 AAACCATTGCTACTGAGTAGTGGCGCGATGGCCACATAAAACTTGAAAAAGCTGACTTACTA   306

Qy   241 AAGGCTATTCAAGAACAACAATTGATCGCTAACGTCCACAGTAACGACTACTTTGAGGTC   300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 AAGGCTATTCAAGAACAACAATTGATCGCTAACGTCCACAGTAACGACTACTTTGAGGTC   366

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAACCATTACTGATCGAACGGAAT   360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAACCATTACTGATCGAACGGAAT   426

Qy   361 AAAGATGGTTCGGTAACCTTGCGACCCAACCGTCCAAGAATTTTGCTAAGCGGACAT   420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 AAAGATGGTTCGGTAACCTTGCGACCCAACCGTCCAAGAATTTTGCTAAGCGGACAT   486

Qy   421 GTGCGCGTTAGACCATATAAAGAAAACAATTACAAGCGAAATCGTTGATGTG   480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 GTGCGCGTTAGACCATATAAAGAAAACAATTACAAGCGAAATCGTTGATGTG   546

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCACAGTCTCAAA   540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCACAGTCTCAAA   606

Qy   541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA   600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA   666

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAACGGCAAGGTCTCACTTTGCTGAC   360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAACGGCAAGGTCTCACTTTGCTGAC   366

Qy   361 AAAGATGGTTCGGTAACCTTGCGACCCAACCGTCCAAGAATTTTGCTAAGCGGACAT   420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 AAAGATGGTTCGGTAACCTTGCGACCCAACCGTCCAAGAATTTTGCTAAGCGGACAT   426

Qy   421 GTGCGCGTTAGACCATATAAAGAAAACAATTACAAGCGAAATCGTTGATGTG   480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 GTGCGCGTTAGACCATATAAAGAAAACAATTACAAGCGAAATCGTTGATGTG   486

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCACAGTCTCAAA   540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCACAGTCTCAAA   540

Qy   541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA   600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA   606

Qy   601 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC   660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC   666

Qy   661 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG   720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG   726

Qy   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG   780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG   786

Qy   781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCTCCTAAAAAGGG   840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   787 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCTCCTAAAAAGGG   846

Qy   841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGTTCACCATCAAATACGTT   900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   847 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGTTCACCATCAAATACGTT   906

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGCGGAGCAGCTCTTAACAGCTAGGGAACGTAAC   960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   907 GATGTCAACACCAACGAATTGCTAAAAAGCGGAGCAGCTCTTAACAGCTAGGGAACGTAAC   966

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC   1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   967 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC   1026

Qy   1021 GATGCTTTGTATTATGGACTATACCTTAACTGAAAAGTAGAGGAATAATCACGATGAC   1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1027 GATGCTTTGTATTATGGACTATACCTTAACTGAAAAGTAGAGGAATAATCACGATGAC   1086

Qy   1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT   1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1087 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT   1146

Qy   1141 CATTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACAGAAGTTTACAGCTACCTG   1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1147 CATTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACAGAAGTTTACAGCTACCTG   1206

Qy   1201 CGTTATACAGGGACACCTTATACCTGATAACCCTAACGACACAATAA   1245
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1207 CGTTATACAGGGACACCTTATACCTGATAACCCTAACGACACAATAA   1251

RESULT 4
A20009
LOCUS   A20009   1317 bp   DNA   linear   PAT 14-JUL-1995
DEFINITION SEQ ID NO: 17; OmpAL fused to a mature streptokinase gene.
ACCESSION   A20009
VERSION   A20009.1   GI:1247841
KEYWORDS   .

```
Qy    61  AGCGTTGCTGGTACTGTTGAGGGGACGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127  AGCGTTGCTGGTACTGTTGAGGGGACGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT   186

Qy   121  GACCTAACATCACGACCTGCTCATGAGAGGAAAGACAGAGAGCAAGGCTTAAGTCCAAAATCA   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187  GACCTAACATCACGACCTGCTCATGAGAGGAAAGACAGAGAGCAAGGCTTAAGTCCAAAATCA   246

Qy   181  AAACCATTTGCTACTGATAGTGGCGGATGCCACATAAAAATCTGGTCTG   240
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   247  AAACCATTTGCTACTGATAGTGGCGGATGCCACATAAAAATCTGGTCTG   306

Qy   241  AAGGCTATTCAAGAACAATGATGCTACGTCACAGTAACGACGACTACTTTGAGGTC   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307  AAGGCTATTCAAGAACAATGATGCTACGTCACAGTAACGACGACTACTTTGAGGTC   366

Qy   301  ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACTGTTCACCATCAAATACGTT   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367  ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACTGTTCACCATCAAATACGTT   426

Qy   361  AAAGATGGTTCGGTAACCTTGCGACCCAACCGTCCAAGAATTTTGTCAAGCGGACAT   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427  AAAGATGGTTCGGTAACCTTGCGACCCAACCGTCCAAGAATTTTGTCAAGCGGACAT   486

Qy   421  GTGCGCGTTAGACCATATAAAGAAGAAAAACAATACAAAATCAAGCGAAATCTGTTGATGTG   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487  GTGCGCGTTAGACCATATAAAGAAGAAAAACAATACAAAATCAAGCGAAATCTGTTGATGTG   546

Qy   481  GAATATACTGTACAGTTTACTCCCTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547  GAATATACTGTACAGTTTACTCCCTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA   606

Qy   541  GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607  GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA   666

Qy   601  GCTCAAGCACACAAAGCATTTAAACAAAAACCCATCCAGGCTATACGACCATTATGAACGTGAC   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667  GCTCAAGCACACAAAGCATTTAAACAAAAACCCATCCAGGCTATACGACCATTATGAACGTGAC   726

Qy   661  TCCTCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727  TCCTCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG   786

Qy   721  TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAAATCTGGTCTG   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   787  TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAAATCTGGTCTG   846

Qy   781  AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTGTCCTTAAAAAGGG   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   847  AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTGTCCTTAAAAAGGG   906

Qy   841  GAAAAGCCGTATGATCCCTTGATCGCAGTCAGTCTTAACAGTTAGCAGTCTAAAACTACGTT   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   907  GAAAAGCCGTATGATCCCTTGATCGCAGTCAGTCTTAACAGTTAGCAGTCTAAAACTACGTT   966

Qy   901  GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTAAAAAGCGAGCTCTTAACAGCTAGCGAACGTAAC   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   967  GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTAAAAAGCGAGCTCTTAACAGCTAGCGAACGTAAC   1026

Qy   961  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTCACAACAATCTC   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1027  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTCACAACAATCTC   1086

Qy  1021  GATGCGTTTTGGTATTATGACTATACCTTAACTGGGCAAGGCAGAAAGTAGAGGATAATCACGATGAC   1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1087  GATGCGTTTTGGTATTATGACTATACCTTAACTGGGCAAGGCAGAAAGTAGAGGATAATCACGATGAC   1146

Qy  1081  ACCAACCGTATCATAACCGTTCAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1147  ACCAACCGTATCATAACCGTTCAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   1206
```

```
Qy  601  GCTCAAGCACACAAAGCATTTAAACAAAAACCCACCCAGGCTATACGACCTTATGAACGTGAC   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  667  GCTCAAGCACACAAAGCATTTAAACAAAAACCCACCCAGGCTATACGACCTTATGAACGTGAC   726

Qy  661  TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCGTACGATCAAGAG   720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  727  TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCGTACGATCAAGAG   786

Qy  721  TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAAATCTGGTCTG   780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  787  TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAAATCTGGTCTG   846

Qy  781  AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAAGTAGAGGATAATCACGATGAC   840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  847  AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAAGTAGAGGATAATCACGATGAC   906

Qy  841  GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTAAAAAGCGAGCTCTTAACAGCTAGCGAACGTAAC   900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  907  GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTAAAAAGCGAGCTCTTAACAGCTAGCGAACGTAAC   966

Qy  901  TTAGACTTCAGAGATTTGGTATTATGACTATACCTTAACTGGAAAAGTAGAGAAGTAGAGAAGTGCTAGCTAT   960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  967  TTAGACTTCAGAGATTTGGTATTATGACTATACCTTAACTGGAAAAGTAGAGAAGTAGAGAAGTGCTAGCTAT   1026

Qy  961  ACCAACCGTATCATAACCGTTATATGGGCAAGGCAGACCCGAAGGAGAGAATGCTAGCTAT   1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1027  ACCAACCGTATCATAACCGTTATATGGGCAAGGCAGACCCGAAGGAGAGAATGCTAGCTAT   1086

Qy 1021  CATTTAGCCTATGATAAGATCGTTATACCGAAGAAGAAGCGAGAAGTTTACAGCTACCTG   1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1087  CATTTAGCCTATGATAAGATCGTTATACCGAAGAAGAAGCGAGAAGTTTACAGCTACCTG   1146

Qy 1081  CGTTATACAGGGACACACCTATACCTGATAACCCTAACGACAAAATAA   1245
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1147  CGTTATACAGGGACACACCTATACCTGATAACCCTAACGACAAAATAA   1311
```

```
RESULT 5
I13197
LOCUS        I13197                   1317 bp      DNA        linear       PAT  26-JUL-1995
DEFINITION   Sequence 18 from patent US 5434073.
ACCESSION    I13197
VERSION      I13197.1      GI:910545
KEYWORDS     .
SOURCE       Unknown.
ORGANISM     Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 1317)
AUTHORS      Dawson,K., Hunter,M.G. and Czaplewski,L.G.
TITLE        Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL      Patent: US 5434073-A 18 18-JUL-1995;
FEATURES             Location/Qualifiers
             source          1..1317
                             /organism="unknown"
                             /mol_type="unassigned DNA"
ORIGIN

Query Match      99.7%;    Score 1241.8;    DB 6;    Length 1317;
Best Local Similarity   99.8%;    Pred. No. 3.2e-296;
Matches 1243;    Conservative   0;    Mismatches   0;    Indels   2;    Gaps   0;

Qy    1  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   126
```

Qy   385 ATTGATTTGCAAGCGATGCAACCATTACTGATCGATGAAGCGCAAGGTCTACTTTGCTGAC 444
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATTGATTTGCAAGCGATGCAACCATTACTGATCGATGAAGCGCAAGGTCTACTTTGCTGAC 420

Qy   445 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAGATTTTTGCTAAGCGGACAT 504
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAGATTTTTGCTAAGCGGACAT 480

Qy   505 GTGCGCGGTTAGACGTTAGACCTATACCTGATAACCCTAACGGACAAATAA 564
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GTGCGCGGTTAGACGTTAGACCTATACCTGATAACCCTAACGGACAAATAA 540

Qy   565 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA 624
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAA 600

Qy   625 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 684
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 660

Qy   685 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC 744
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GCTCAAGCACAAAGCATTTAAACAAAACCATCCAGGCTATACGATTTATGAACGTGAC 720

Qy   745 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 804
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 780

Qy   805 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 864
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 840

Qy   865 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCCTTAAAAAAGGG 924
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCCTTAAAAAAGGG 900

Qy   925 GAAAAGCCGTATGATCCCTTGATCGGCAGTCACTGAAACTGTTCACCATCAAATACGTT 984
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAAAAGCCGTATGATCCCTTGATCGGCAGTCACTGAAACTGTTCACCATCAAATACGTT 960

Qy   985 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1044
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1020

Qy 1045 TTAGACTTCAGAGATTTATACAGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1104
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 TTAGACTTCAGAGATTTATACAGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1080

Qy 1105 GATGCTTTTGGTATTATGGACTATACCTTAACTGGACTATACCTTAACTGGACGAAGATC 1164
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 GATGCTTTTGGTATTATGGACTATACCTTAACTGGACTATACCTTAACTGGACGAAGATA 1140

Qy 1165 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1224
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1200

Qy 1225 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAAGTTTACAGCTACCTG 1284
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAAGTTTACAGCTACCTG 1245

Qy 1285 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA 1329
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 1285 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA 1329

RESULT 7
I13194
LOCUS   I13194   1335 bp   DNA   linear   PAT 26-JUL-1995
DEFINITION Sequence 14 from patent US 5434073.
ACCESSION   I13194
VERSION   I13194.1 GI:910542
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.

Qy 1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAAGTTTACAGCTACCTG 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1207 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAAGTTTACAGCTACCTG 1266

Qy 1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA 1245
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 1267 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA 1311

RESULT 6
A20006
LOCUS   A20006   1335 bp   DNA   linear   PAT 14-JUL-1995
DEFINITION SEQ ID NO: 14; Streptokinase gene.
ACCESSION   A20006
VERSION   A20006.1 GI:1247579
KEYWORDS   .
SOURCE   Streptococcus dysgalactiae subsp. equisimilis
  ORGANISM Streptococcus dysgalactiae subsp. equisimilis
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE   1 (bases 1 to 1335)
  AUTHORS   .
  TITLE   PROTEINS AND NUCLEIC ACIDS
  JOURNAL   Patent: WO 9109125-A 14 27-JUN-1991;
FEATURES   Location/Qualifiers
     source   1..1335
                     /organism="Streptococcus dysgalactiae subsp. equisimilis"
                     /mol_type="unassigned DNA"
                     /sub_species="equisimilis"
                     /db_xref="taxon:119602"
     CDS   7..1329
                     /codon_start=1
                     /transl_table=11
                     /product="streptokinase"
                     /protein_id="CAA01484.1"
                     /db_xref="GI:124758"
                     /db_xref="REMTREMBL:CAA01484"
                     /translation="MKNYLSFGMFALLFALTFGTVNSVQAIAGPEWLLDRPSVNNSQL
                     VVSVAGTVEGTNQDISLKFFEIDLTSRPAHGGKTEQGLSPKSPFATDSGAMPHKLEK
                     ADLLKAIDEQLIANVHSNDDYFEVIDFASDATTDRNGKVYFADKDGSVTLPTQPVQE
                     FLLSGHVRVRPYKEKPIQNQAKSVDVEYTVQFTLNPDDFRPGLKDTKLLKTLAIGD
                     TITSQELLAQAQSIINKTHPGYTIYERDSSIVTHDNDIFRTILPMDQEFTYHVKNREQ
                     AYRINKKSGLNEEINNTDLISEKYYVLKKGEKPYDPFDRSHLKLFTIKYVDVNTNELL
                     KSEQLLTASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDDTNRIIT
                     VYMGKRPEGNASYHLAYDKDRYTEBEREVYSLRYTGTPIPDNPNDK"
ORIGIN

Query Match   99.7%; Score 1241.8; DB 6; Length 1335;
Best Local Similarity   99.8%; Pred. No. 3.2e-296;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 ATTGCTGGACCTGAGTGGCTGCTCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 ATTGCTGGACCTGAGTGGCTGCTCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 144

Qy   61 AGCGGTTGCTGGTGACTCGTTGAGGGGACGAATGAATCAAGACATTAGTTCTTAAATTTTGAAATT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145 AGCGGTTGCTGGTGACTCGTTGAGGGGACGAATGAATCAAGACATTAGTTCTTAAATTTTGAAATT 204

Qy 121 GACCTAACATCACGACCTGCTCATGGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 205 GACCTAACATCACGACCTGCTCATGGAGAGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 264

Qy 181 AAACCATTTGCTACTGATAGTTGCGCGATGCCGCGCCATAAAACTTGAAAAAGCTGACTTACTA 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 265 AAACCATTTGCTACTGATAGTTGCGCGATGCCGCGCCATAAAACTTGAAAAAGCTGACTTACTA 324

Qy 241 AAGGCTATTCAAGAACAATTGATCGCTAACGTTCCCAGTAACGACGACTACTTTGGAGGTC 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 325 AAGGCTATTCAAGAACAATTGATCGCTAACGTTCCCAGTAACGACGACTACTTTGGAGGTC 384

Qy 301 ATTGATTTGCAAGCGATGCAACCATTACTGATCGATGAAGCGCAAGGTCTACTTTGCTGAC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

REFERENCE   1 (bases 1 to 1335)
AUTHORS   Dawson,K., Hunter,M.G. and Czaplewski,L.G.
TITLE   Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL   Patent: US 5434073-A 14 18-JUL-1995;
FEATURES   Location/Qualifiers
  source   1..1335
             /organism="unknown"
             /mol_type="unassigned DNA"
ORIGIN

Query Match   99.7%; Score 1241.8; DB 6; Length 1335;
Best Local Similarity   99.8%; Pred. No. 3.2e-296;
Matches 1243; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy      1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACGCCAATTAGTTGTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 144

Qy     61 AGGGTTGCTGTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 AGGGTTGCTGTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 204

Qy    121 GACCTAACATCACGACCTGCTCATGGTAACCTGCCGACCCAACCTGTCTAAGGCTTACTA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 GACCTAACATCACGACCTGCTCATGGTAACCTGCCGACCCAACCTGTCTAAGGCTTACTA 264

Qy    181 AAACCATTTGCTACTGATAGTGGCGCGATGGCCGATGCCACATAAACTTGAAAAAGCTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 AAACCATTTGCTACTGATAGTGGCGCGATGGCCGATGCCACATAAACTTGAAAAAGCTAC 324

Qy    241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 384

Qy    301 ATTGATTTTGCAAGCGATGCAACCATTACTGATGGAAACGGCAAGGTCTCACTTGCTGAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 ATTGATTTTGCAAGCGATGCAACCATTACTGATGGAAACGGCAAGGTCTCACTTGCTGAC 444

Qy    361 AAAGATGGTTCGGTAACCTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    445 AAAGATGGTTCGGTAACCTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT 504

Qy    421 GTGCGGCGTTAGACAGTATAAAGAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 GTGCGGCGTTAGACAGTATAAAGAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 564

Qy    481 GAATATACTGTACAGTTACTCCCCTTAAACCCTGATGACGATTTCAGACGACCAGGTCTCAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    565 GAATATACTGTACAGTTACTCCCCTTAAACCCTGATGACGATTTCAGACGACCAGGTCTCAAA 624

Qy    541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACGACCATTACTCCAAGAATTACTA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACGACCATTACTCCAAGAATTACTA 684

Qy    601 GCTCAAGCACAAAGCATTTTAAACAAAACCACCCAGGCTATACGATTTATGACGTGAC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    685 GCTCAAGCACAAAGCATTTTAAACAAAACCACCCAGGCTATACGATTTATGACGTGAC 744

Qy    661 TCCTCAATGTCACTCATGCATGAACAAGCCTATGAACAAGCTTATGACAATTTCCGTACGTTACCATGGATCAAGAG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    745 TCCTCAATGTCACTCATGCATGAACAAGCCTATGAACAAGCTTATGACAATTTCCGTACGTTACCATGGATCAAGAG 804

Qy    721 TTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCGGTCTG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    805 TTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCGGTCTG 864

Qy    781 AATGAAGAAATAAACAACACTGACCTGACCTGATCTCTGAGAATATTACGTCCTTAAAAAGGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    865 AATGAAGAAATAAACAACACTGACCTGACCTGATCTCTGAGAATATTACGTCCTTAAAAAGGG 924

Qy    841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTGAACTGTCAATCACCATCAATACGTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    925 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTGAACTGTCACCATCAATACGTT 984
```

RESULT 8
A20027
LOCUS   A20027   1458 bp   DNA   linear   PAT 14-JUL-1995
DEFINITION   SEQ ID NO: 35; Synthetic nucleotide sequence for
             Hirudin-streptokinase fusion protein.
ACCESSION   A20027
VERSION   A20027.1 GI:1247862
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
             1 (bases 1 to 1458)
REFERENCE
  AUTHORS
  TITLE   PROTEINS AND NUCLEIC ACIDS
  JOURNAL   Patent: WO 9109125-A 35 27-JUN-1991;
FEATURES   Location/Qualifiers
  source   1..1458
             /organism="synthetic construct"
             /mol_type="unassigned DNA"
             /db_xref="taxon:32630"
ORIGIN

Query Match   99.7%; Score 1241.8; DB 6; Length 1458;
Best Local Similarity   99.8%; Pred. No. 3.2e-296;
Matches 1243; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy      1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACGCCAATTAGTTGTT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 267

Qy     61 AGCGGTTGCTGTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 AGCGGTTGCTGTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 327

Qy    121 GACCTAACATCACGACCTGCTCATGGTGACACCTGCTCATGGTGACACCTGCCAAAATCA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 GACCTAACATCACGACCTGCTCATGGTGACACCTGCTCATGGTGACACCTGCCAAAATCA 387

Qy    181 AAACCATTTGCTACTGATAGTGGCCGATGCCGATGCCGATGCCACATAAACTTGAAAAGCTACTA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388 AAACCATTTGCTACTGATAGTGGCCGATGCCGATGCCGATGCCACATAAACTTGAAAAAGCTACTA 447

Qy    241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 507

Qy    301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGCAAGGTCTCACTTGCTGAC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy    901 GATGTCAACACCAACCAACGAATTGCTAAAAAGGCAGCAGCTTCTTAACAGTCAGCGAACGTAAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    985 GATGTCAACACCAACCAACGAATTGCTAAAAAGGCAGCAGCTTCTTAACAGTCAGCGAACGTAAC 1044

Qy    961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACACAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1045 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACACAATCTC 1104

Qy   1021 GATGCTTTGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105 GATGCTTTGTATTATGGACTATACCTTAACTGGAAAAGTAGAAGATAATCACGATGAC 1164

Qy   1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1224

Qy   1141 CATTAGCCATGATAAAGATCGTTATACCGAAGAGACAAGCTTGTTACAGCTACCTG 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225 CATTAGCCATGATAAAGATCGTTATACCGAAGAGACAAGCTTGTTACAGCTACCTG 1284

Qy   1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACCACAAATAA 1245
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1285 CGTTATACAGGGACACCTATACCTGATAACCCTAACCACAAATAA 1329
```

REFERENCE   1 (bases 1 to 1458)
AUTHORS   Dawson,K., Hunter,M.G. and Czaplewski,L.G.
TITLE   Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL   Patent: US 5434073-A 42 18-JUL-1995;
FEATURES   Location/Qualifiers
source   1..1458
             /organism="unknown"
             /mol_type="unassigned DNA"
ORIGIN

Query Match   99.7%;   Score 1241.8; DB 6; Length 1458;
Best Local Similarity 99.8%;   Pred. No. 3.2e-296;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 ATTGCTGGACCTGAGTGCTGCTAGACCGTCGTAGACGTTCATCTGTCAACAACAGCAACTAGTTGTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCGTAGACGTTCATCTGTCAACAACAGCAACTAGTTGTT 267

Qy     61 AGCGTTGCTCGGTACGTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 AGCGTTGCTCGGTACGTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 327

Qy    121 GACCTAACAATCACAGACCTGCTCATGAGGAAAGACAGAGCAAGCTTAAGTCCAAAATCA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 GACCTAACAATCACAGACCTGCTCATGAGGAAAGACAGAGCAAGCTTAAGTCCAAAATCA 387

Qy    181 AAACCATTTGCTACTGATAGTGGCGCGATGCCTAACGTCCACATAAACTTGAAAAAGCTGACTTACTA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388 AAACCATTTGCTACTGATAGTGGCGCGATGCCTAACGTCCACATAAACTTGAAAAAGCTGACTTACTA 447

Qy    241 AAGGCTATTCAAGAACAATTGATCGCTAACATTACTACTCCGACCCAACGTCCACAGTAACGTCGAAACGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448 AAGGCTATTCAAGAACAATTGATCGCTAACATTACTACTCCGACCCAACGTCCACAGTAACGTCGAAACGTC 507

Qy    301 ATTGATTTTGCAAGCGATGCAACATTACTACTCGAAACTGTTCACTTGCTGAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    508 ATTGATTTTGCAAGCGATGCAACATTACTACTCGAAACTGTTCACTTGCTGAC 567

Qy    361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGGACAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    568 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGGACAT 627

Qy    421 GTGCGCGTTAGACCGTTAGACTATACGACTAAGGCTAAACTGGAAATCGTTGATGTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    628 GTGCGCGTTAGACCGTTAGACTATACGACTAAGGCTAAACTGGAAATCGTTGATGTG 687

Qy    481 GAATATACTGTACAGTTACTCCCTAAACCCTAAAACCTGATGACGATTTCAGACCAGGTCTCAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    688 GAATATACTGTACAGTTACTCCCTAAACCCTAAAACCTGATGACGATTTCAGACCAGGTCTCAAA 747

Qy    541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGGTGACACCATCCAAGAATTACTA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    748 GATACTAAGCTATTGAAAACACTAGCTATCGGTGGTGACACCATCCAAGAATTACTA 807

Qy    601 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGTATCCTG 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    808 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGTATCCTG 867

Qy    661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGACGATTTTACCAATGGATCAAGAG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    868 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGACGATTTTACCAATGGATCAAGAG 927

Qy    721 TTTACTTACCATGTCAAAATCGGGACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    928 TTTACTTACCATGTCAAAATCGGGACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 987

Qy    781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    988 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1047

Qy    841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    508 ATTGATTTTGCAAGCGATGCAACATTACTACTCGAAACTGTTCACTTGCTGAC 567

Qy    361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGGACAT 420

Db    568 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGGACAT 627

Qy    421 GTGCGCGTTAGACCGTTAGACTATACGACTAAGGCTAAACTGGAAATCGTTGATGTG 480

Db    628 GTGCGCGTTAGACCGTTAGACTATACGACTAAGGCTAAACTGGAAATCGTTGATGTG 687

Qy    481 GAATATACTGTACAGTTACTCCCTAAACCCTAAAACCTGATGACGATTTCAGACCAGGTCTCAAA 540

Db    688 GAATATACTGTACAGTTACTCCCTAAACCCTAAAACCTGATGACGATTTCAGACCAGGTCTCAAA 747

Qy    541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGGTGACACCATCCAAGAATTACTA 600

Db    748 GATACTAAGCTATTGAAAACACTAGCTATCGGTGGTGACACCATCCAAGAATTACTA 807

Qy    601 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGTATCCTG 660

Db    808 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGTATCCTG 867

Qy    661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGACGATTTTACCAATGGATCAAGAG 720

Db    868 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGACGATTTTACCAATGGATCAAGAG 927

Qy    721 TTTACTTACCATGTCAAAATCGGGACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780

Db    928 TTTACTTACCATGTCAAAATCGGGACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 987

Qy    781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840

Db    988 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1047

Qy    841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900

Db   1048 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 1107

Qy    901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 960

Db   1108 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1167

Qy    961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTCTACAACAATCTC 1020

Db   1168 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTCTACAACAATCTC 1227

Qy   1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAGTAGAGGATAATCACGATGAC 1080

Db   1228 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAGTAGAGGATAATCACGATGAC 1287

Qy   1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAAATGCTAGCTAT 1140

Db   1288 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAAATGCTAGCTAT 1347

Qy   1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200

Db   1348 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1407

Qy   1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACCGACACAATAA 1245

Db   1408 CGTTATACAGGGACACCTATACCTGATAACCCTAACCGACACAATAA 1452
```

RESULT 9
I13215
LOCUS   I13215   1458 bp   DNA   linear   PAT 26-JUL-1995
DEFINITION Sequence 42 from patent US 5434073.
ACCESSION   I13215
VERSION   I13215.1   GI:910563
KEYWORDS   .
SOURCE   Unknown.
ORGANISM   Unknown.

```
Db    61 AGCGGTTGCTCGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   322 AGCGTTGCTCGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 381

Db   121 GACCTAACATCACGACCTGCTCATGAGAGGCTAAAACACTCTTAAGTCCAAAATCA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   382 GACCTAACATCACGACCTGCTCATGAGAGGCTAAAACACTCTTAAGTCCAAAATCA 441

Db   181 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAAACTGAAAAAGCTGACTACTA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   442 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAAACTGAAAAAGCTGACTACTA 501

Db   241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTGGAGTC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   502 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTGGAGTC 561

Db   301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGTCTCACTTGCTGAC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   562 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGTCTCACTTGCTGAC 621

Db   361 AAAGATGGTTCGGTAACCTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACACAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   622 AAAGATGGTTCGGTAACCTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACACAT 681

Db   421 GTGCGCGTTAGACCATATAAAGAAAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   682 GTGCGCGTTAGACCATATAAAGAAAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 741

Db   481 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACCATTTCGGTACGATTTCAGACCAGGTCTCAAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   742 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACCATTTCGGTACGATTTCAGACCAGGTCTCAAA 801

Db   541 GATACTAAGCTATTGAAAACACTAGCTATGCGGTATCGGTGACACCATCACATCTCAAGAATTACTA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   802 GATACTAAGCTATTGAAAACACTAGCTATGCGGTATCGGTGACACCATCACATCTCAAGAATTACTA 861

Db   601 GCTCAAGCACAAAGCATTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   862 GCTCAAGCACAAAGCATTTAAACAAAACCCATCCAGGCTATACGATTTATGAACGTGAC 921

Db   661 TCCTCAATCGTCACTCATGACAATGACACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   922 TCCTCAATCGTCACTCATGACAATGACACATTTTCCGTACGATTTTACCAATGGATCAAGAG 981

Db   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGGAACAAGCTTATGAGATCAATAAAAAAATCTCGGTCTG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   982 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAAATCTCGGTCTG 1041

Db   781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCTTAAAAAAGGAG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1042 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCTTAAAAAAGGG 1101

Db   841 GAAAGCCGTATGATCCCTTGATCGCAGTCACTTGAAACTTGAAACTGTTCACCATCAAATACGTT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1102 GAAAGCCGTATGATCCCTTGATCGCAGTCACTTGAAACTTGAAACTGTTCACCATCAAATACGTT 1161

Db   901 GATGTCAACACCAACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTGCGAACGTAAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1162 GATGTCAACACCAACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTGCGAACGTAAC 1221

Db   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1222 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1281

Db  1021 GATGCTTTTGGTATTATGGACTATACCTAACTGACGAAAAGTAGACGGATAATCACGATGAC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1282 GATGCTTTTGGTATTATGGACTATACCTAACTGACGAAAAGTAGAAGTAAAATCACGATGAC 1341

Qy  1081 ACCAACCGTATCATAACCGTTCATATGGGCAAGCGACACCCGAAGGAGAGAATGCTAGCTAT 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1342 ACCAACCGTATCATAACCGTTCATATGGGCAAGCGACACCCGAAGGAGAGAATGCTAGCTAT 1401

Qy  1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1202 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
```

```
Db  1048 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 1107
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901 GATGTCAACACCAACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 960

Db  1108 GATGTCAACACCAACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1167
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAATGGCTAAACTACTCTACAACAATCTC 1020

Db  1168 TTAGACTTCAGAGATTTATACGATCCTCGTGATAATGGCTAAACTACTCTACAACAATCTC 1227
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1021 GATGCTTTTGGTATTATGGACTATACCGTTTATGTGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1080

Db  1228 GATGCTTTTGGTATTATGGACTATACCGTTTATGTGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1287
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1081 ACCAACCGTATCATAACCGTTTATGTGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1140

Db  1288 ACCAACCGTATCATAACCGTTTATGTGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1347
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAAGTTTACAGCTACCTG 1200

Db  1348 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAAGTTTACAGCTACCTG 1407

Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
          |||||||||||||||||||||||||||||||||||||||||||||
Db  1408 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1452
```

RESULT 10

```
A20016
LOCUS       A20016                  1512 bp    DNA     linear   PAT 14-JUL-1995
DEFINITION  Nucleotide sequence for streptokinase fused to yeast
            alpha-factor.
ACCESSION   A20016
VERSION     A20016.1  GI:1247850
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1512)
  AUTHORS   PROTEINS AND NUCLEIC ACIDS
  TITLE     
  JOURNAL   Patent: WO 9109125-A 24 27-JUN-1991;
FEATURES             Location/Qualifiers
     source          1..1512
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
     CDS             7..1506
                     /codon_start=1
                     /transl_table=11
                     /product="streptokinase fused to yeast alpha-factor"
                     /protein_id="CAA01487.1"
                     /db_xref="GI:1247851"
                     /db_xref="REMTREMBL:CAA01487"
                     /translation="MRFPSIFTAVLFAASSALAAPVNTTEDETAQIPAEAVIGYLDL
                     EGDFDVAVLPFSNSTNNGLLFINTTIASIAAKEEGVSLDKRIAGPEWLLDRPSVNNSQ
                     LVVSVAGTVEGTNQDISLKFFEIDLTSRPAHGGKTEQGLSPKSKPFATDSGAMPHKLE
                     KADLLKAIQEQIANVHSNDDYFEVIDPASDATITDRNGKYFADKDGSVTLPTQPVQ
                     EFLLSGHVRRVRPYKEKPIQNQAKSVDEYTVQFTPLNPDDDFRPGLKDTKLLKTLAIG
                     DTITSQELLAQAQSILNKTHPGYTIYERDSSIVTHNDIFRTILPMDQEFYHVKNRE
                     QAYEINKKSGLNEEINNTDLISEKYVLKKGEKPYDPFDRSHLKLFTIKVDVNTNEL
                     LKSEQLTASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDDTNRII
                     TVYMGKRPEGENASYHLAYDKDRYTEEREVSYLRYTGTPIPDNPNDK"
ORIGIN
```

```
Query Match       99.7%;  Score 1241.8;  DB 6;  Length 1512;
Best Local Similarity 99.8%;  Pred. No. 3.2e-296;
Matches 1243;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;
```

```
Qy     1 ATTGCTGGACCTGAGTCGCTGCTAGACCGTCCATCTGCTCAACAACAGCCAATTAGTTGTT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262 ATTGCTGGACCTGAGTCGCTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 321

Qy  1141 CATTTAGCCTATGATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
```

```
Db    1402  CATTTAGCCTATGATAAAGATCGTTATACCGATAACCCTAACGACACAAATAA 1245
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     862  GCTCAAGCACAAAGCATTTAAACAAAACCCATCCAGCTATACGATTTATGAACGTGAC 921

Qy    1201  CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA 1245
            |||||||||||||||||||||||||||||||||||||||||||||||
Qy     661  TCCTCAATCGTCACTCACTGACAATGACATTTTCCGTACGACATTTTACCAATGGATCAAGAG 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1462  CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA 1506
Db     922  TCCTCAATCGTCACTCACTGACAATGACATTTTCCGTACGACATTTTACCAATGGATCAAGAG 981

RESULT 11                                              Qy     721  TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
I13204                                                             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
LOCUS       I13204        1512 bp  DNA     linear  PAT 26-JUL-1995  Db     982  TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 1041
DEFINITION  Sequence 27 from patent US 5434073.
ACCESSION   I13204                                   Qy     781  AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
VERSION     I13204.1  GI:910552                                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
KEYWORDS    .                                        Db    1042  AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 1101
SOURCE      Unknown.
  ORGANISM  Unknown.                                  Qy     841  GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900
            Unclassified.                                         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
REFERENCE   1  (bases 1 to 1512)                      Db    1102  GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 1161
  AUTHORS   Dawson,K., Hunter,M.G. and Czaplewski,L.G.
  TITLE     Fibrinolytic and anti-thrombotic cleavable dimers   Qy     901  GATGTCAACACCAACGAATTGCTAAAAAGCAGCAGCTCTTAACAGCTAGCGGAACGTAAC 960
  JOURNAL   Patent: US 5434073-A 27 18-JUL-1995;                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
FEATURES             Location/Qualifiers               Db    1162  GATGTCAACACCAACGAATTGCTAAAAAGCAGCAGCTCTTAACAGCTAGCGGAACGTAAC 1221
     source          1..1512
                     /organism="unknown"              Qy     961  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACACAATCTC 1020
                     /mol_type="unassigned DNA"                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ORIGIN                                                Db    1222  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACACAATCTC 1281

    Query Match      99.7%;  Score 1241.8; DB 6; Length 1512;    Qy    1021  GATGCTTTTGGTATTATGGACTATACCTAACTGGAAAAGTAGAGGATAATACGATGAC 1080
    Best Local Similarity   99.8%;  Pred. No. 3.2e-296;                      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Matches 1243;  Conservative  0;  Mismatches  2;  Indels  2;  Gaps  0;  Db    1282  GATGCTTTTGGTATTATGGACTATACCTAACTGGAAAAGTAGAAGATAATCGATGAC 1341

Qy       1  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60   Qy    1081  ACCAACGTATCATAACCGTTTATATGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     262  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 321   Db    1342  ACCAACGTATCATAACCGTTTATATGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1401

Qy      61  AGCGTTGCTGGTACTGTTGAGGGACGCAACCATTAGTCTTAAATTTTTTGAAATT 120    Qy    1141  CATTTAGCCTATGATAAAGATCGTTATACCGATAACCCTAACGACACAAATAA 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||                     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     322  AGCGTTGCTGGTACTGTTGAGGGACGCAACCATTAGTCTTAAATTTTTTGAAATT 381    Db    1402  CATTTAGCCTATGATAAAGATCGTTATACCGATAACCCTAACGACACAAATAA 1461

Qy     121  GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180   Qy    1201  CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA 1245
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||               |||||||||||||||||||||||||||||||||||||||||||||||
Db     382  GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 441   Db    1462  CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA 1506

Qy     181  AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTACTA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||   RESULT 12
Db     442  AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTACTA 501   I13209
                                                                         LOCUS       I13209        2589 bp  DNA     linear  PAT 26-JUL-1995
Qy     241  AAGGCTATTCAAGAACAATTGATCGTCAACGTTCCAGTAACGACGACTACTTTGAGGTC 300   DEFINITION  Sequence 34 from patent US 5434073.
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||   ACCESSION   I13209
Db     502  AAGGCTATTCAAGAACAATTGATCGTCAACGTTCCAGTAACGACGACTACTTTGAGGTC 561   VERSION     I13209.1  GI:910557
                                                                         KEYWORDS    .
Qy     301  ATTGATTTTGCAAGCGATGCAACCATTACTGATGAAACGGCAAGGTCTCACTTTGCTGAC 360   SOURCE      Unknown.
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||     ORGANISM  Unknown.
Db     562  ATTGATTTTGCAAGCGATGCAACCATTACTGATGAAACGGCAAGGTCTCACTTTGCTGAC 621             Unclassified.
                                                                         REFERENCE   1  (bases 1 to 2589)
Qy     361  AAAGATGGTTCGGTAACCTTGCCAGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420    AUTHORS   Dawson,K., Hunter,M.G. and Czaplewski,L.G.
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||     TITLE     Fibrinolytic and anti-thrombotic cleavable dimers
Db     622  AAAGATGGTTCGGTAACCTTGCCAGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 681    JOURNAL   Patent: US 5434073-A 34 18-JUL-1995;
                                                                         FEATURES             Location/Qualifiers
Qy     421  GTGCGCGTTAGACCTTACTCCTTAAACAAATACAAAATCAAGCGAAATCTGTTGATGTG 480        source          1..2589
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||                         /organism="unknown"
Db     682  GTGCGCGTTAGACCTTACTCCTTAAACAAATACAAAATCAAGCGAAATCTGTTGATGTG 741    ORIGIN                   /mol_type="unassigned DNA"

Qy     481  GAATATACTGTACAGTTTACTCCGTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA 540     Query Match      99.7%;  Score 1241.8; DB 6; Length 2589;
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||     Best Local Similarity   99.8%;  Pred. No. 3.2e-296;
Db     742  GAATATACTGTACAGTTTACTCCGTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA 801     Matches 1243;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy     541  GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 600   Qy       1  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     802  GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 861   Db    1339  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 1398

Qy     601  GCTCAAGCACAAAGCATTTAAACAAAACCCACCCAGCTATACGATTTATGAACGTGAC 660   Qy      61  AGCGTTGCTGGTACTGTTGAGGGACGCAACCATTAGTCTTAAATTTTTTGAAATT 120
```

Db   2479 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTACAGCTACCTG 2538

Qy   1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
          |||||||||||||||||||||||||||||||||||||||||||||
Db   2539 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 2583

RESULT 13
A20030
LOCUS   A20030   1467 bp   DNA   linear   PAT 14-JUL-1995
DEFINITION SEQ ID NO: 38; Synthetic nucleotide sequence for
            streptokinase-hirudin fusion protein.
ACCESSION   A20030
VERSION   A20030.1 GI:1247865
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
            1 (bases 1 to 1467)
REFERENCE   .
  AUTHORS   PROTEINS AND NUCLEIC ACIDS
  TITLE   .
  JOURNAL   Patent: WO 9109125-A 38 27-JUN-1991;
FEATURES   Location/Qualifiers
     source   1..1467
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

Query Match   99.5%; Score 1238.8; DB 6; Length 1467;
Best Local Similarity 99.8%; Pred. No. 1.8e-295;
Matches 1240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT 60

Qy   61 AGCGTTGCTGTGGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTGAAATT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGCGTTGCTGTGGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTGAAATT 120

```
Qy   61 AGGGTTGCTGGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGCGTTGCTGGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT  120

Qy  121 GACCTAACATCACGACCTGCTCATGACGATTTCCGTACGATTTACCAATGACTCA  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GACCTAACATCACGACCTGCTCATGACGATTTCCGTACGATTTACCAATGACTCA  180

Qy  181 AAACCATTTGCTACTGACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  240
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AAACCATTTGCTACTGACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  240

Qy  241 AAGGCTATTCAAGAACAACAACTGATCGCTAACGTCCAGTAACGACTACTTGAGGTC  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AAGGCTATTCAAGAACAACAACTGATCGCTAACGTCCAGTAACGACTACTTGAGGTC  300

Qy  301 ATTGATTTTGCAAGCGATGCAACCATTACTGACGTTCACCCATCAAATACGTT  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ATTGATTTTGCAAGCGATGCAACCATTACTGACGTTCACCCATCAAATACGTT  360

Qy  361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420

Qy  421 GTGCGCGTTAGACCATATAAAGAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTGCGCGTTAGACCATATAAAGAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG  480

Qy  481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  540

Qy  541 GATACTAAGCTATTGAAAAACACTAGCTGATCGGTGACACCACCATCACATCTCAAGAATTACTA  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GATACTAAGCTATTGAAAAACACTAGCTGATCGGTGACACCACCATCACATCTCAAGAATTACTA  600

Qy  601 GCTCAAGCACAAAGCATTTAAACAAAACCCATCCAAGCCTATACAGATTATGAACGTGAC  660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GCTCAAGCACAAAGCATTTAAACAAAACCCATCCAAGCCTATACAGATTATGAACGTGAC  660

Qy  661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG  720

Qy  721 TTTACTTACCATGTCAAAATCGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TTTACTTACCATGTCAAAATCGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  780

Qy  781 AATGAAGAAATAAACACACTACTGACCTCTGAGAAATATTACGTCCTTAAAAAAGGG  840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AATGAAGAAATAAACACACTACTGACCTCTGAGAAATATTACGTCCTTAAAAAAGGG  840

Qy  841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  900

Qy  901 GATGTCAACACCAACGAATTGCTAAAAAGCAGCTCTTAACAGCTAGCGAACGTAAC  960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 GATGTCAACACCAACGAATTGCTAAAAAGCAGCTCTTAACAGCTAGCGAACGTAAC  960

Qy  961 TTAGACTTCAGAGATTTATACGATCCTGTGATAAGGCTAAACTACTCCACAACAATCTC 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTAGACTTCAGAGATTTATACGATCCTGTGATAAGGCTAAACTACTCCACAACAATCTC 1020

Qy 1021 GATGCTGTTTGGTATTATGGACTATACCTTAACTCGGAGATAATAATCACGATGAC 1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 GATGCTGTTTGGTATTATGGACTATACCTTAACTCGGAGATAATAATCACGATGAC 1080

Qy 1081 ACCAACGTATCATAACCGTTTATGGGCAAGCCACCCGAAGGAGAGAATGCTAGCTAT 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 ACCAACGTATCATAACCGTTTATGGGCAAGCCACCCGAAGGAGAGAATGCTAGCTAT 1140

Qy 1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
```

```
Qy  601 GCTCAAGCACAAAGCATTTAAACAAAACCCACCCAGGCTATACGATTATGAACGTGAC  660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GCTCAAGCACAAAGCATTTAAACAAAACCCACCCAGGCTATACGATTATGAACGTGAC  660

Qy  661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG  720

Qy  721 TTTACTTACCATGTCAAAATCGGGACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TTTACTTACCATGTCAAAATCGGGACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  780

Qy  781 AATGAAGAAATAAACACACTGGGACAAGCAAGCTTATGGCAAGCTTATGGGCAAGCT  840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AATGAAGAAATAAACACACTGGGACAAGCAAGCTTATGGCAAGCTTATGGGCAAGCT  840

Qy  841 GAAAAGCCGTATGATCCCTTTGATCGGCAGTCACTTGAAACTGTTCACCTGAAATACGTT  900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GAAAAGCCGTATGATCCCTTTGATCGGCAGTCACTTGAAACTGTTCACCTGAAATACGTT  900

Qy  901 GATGTCAACACCAACGAATTGCTAAAAAGCTAAAAAGCTAAAAAGCTAAAAAGGG  960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 GATGTCAACACCAACGAATTGCTAAAAAGCTAAAAAGCTAAAAAGCTAAAAAGGG  960

Qy  961 TTAGACTTCAGAGATTTATACGATCCTGTGATAAGGCTAAACTACTCCACAACAATCTC 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTAGACTTCAGAGATTTATACGATCCTGTGATAAGGCTAAACTACTCCACAACAATCTC 1020

Qy 1021 GATGTCTTTGGTATTATGGACTATACCTTAACTGACGTTCACCATCAAATACGTT 1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 GATGTCTTTGGTATTATGGACTATACCTTAACTGACGTTCACCATCAAATACGTT 1080

Qy 1081 ACCAACCGTATCATAACCGTTTATGGGCAAGCCACCCGAAGGAGAGAATGCTAGCTAT 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 ACCAACCGTATCATAACCGTTTATGGGCAAGCCACCCGAAGGAGAGAATGCTAGCTAT 1140

Qy 1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200

Qy 1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA 1242
        ||||||||||||||||||||||||||||||||||||||||||
Db 1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA 1242
```

RESULT 14
I13218
LOCUS   I13218   1467 bp   DNA   linear   PAT   26-JUL-1995
DEFINITION Sequence 46 from patent US 5434073.
ACCESSION   I13218
VERSION   I13218.1 GI:910566
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1467)
  AUTHORS   Dawson,K., Hunter,M.G. and Czaplewski,L.G.
  TITLE   Fibrinolytic and anti-thrombotic cleavable dimers
  JOURNAL   Patent: US 5434073-A 46 18-JUL-1995;
FEATURES   Location/Qualifiers
     source   1..1467
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   99.5%; Score 1238.8; DB 6; Length 1467;
Best Local Similarity 99.8%; Pred. No. 1.8e-295;
Matches 1240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   60

```
Db  1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA 1242
              ||||||||||||||||||||||||||||||||||||||||||
Db  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA 1242

RESULT 15
A20021
LOCUS       A20021      2588 bp    DNA    linear    PAT 14-JUL-1995
DEFINITION  SEQ ID NO: 29; Synthetic nucleotide sequence for
            OmpAL-streptokinase-streptokinase fusion gene.
ACCESSION   A20021
VERSION     A20021.1   GI:1247856
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 2588)
  AUTHORS   .
  TITLE     PROTEINS AND NUCLEIC ACIDS
  JOURNAL   Patent: WO 9109125-A 29 27-JUN-1991;
            1. .2588
FEATURES             Location/Qualifiers
     source          1..2588
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

Query Match     99.5%;   Score 1238.8;   DB 6;   Length 2588;
Best Local Similarity   99.8%;   Pred. No. 1.8e-295;
Matches 1240;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT  126

Qy    61 AGCGGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTGAAATT  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 AGCGGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTGAAATT  186

Qy   121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  246

Qy   181 AAACCATTTGCTACTACTAGTGGCGCCGATGGCCCACATAACTTGAAAAAGTGACTTACTA  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 AAACCATTTGCTACTACTAGTGGCGCCGATGGCCCACATAAACTTGAAAAAGTGACTTACTA  306

Qy   241 AAGGCTATTCAAGAACAATTGATCGTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 AAGGCTATTCAAGAACAATTGATCGTAACGTCCACAGTAACGACGACTACTTTGAGGTC  366

Qy   301 ATTGATTTTGCAAGCGATGCAACATTACTTACTGATCGAAACGGCAAGGTCTCACTTTGCTGAC  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 ATTGATTTTGCAAGCGATGCAACATTACTTACTGATCGAAACGGCAAGGTCTCACTTTGCTGAC  426

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT  486

Qy   421 GTGCGCCGTTAGACCATATAAGAAAAACAATACAAAATCAAGCGAAATCTGTTGATGTG  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 GTGCGCCGTTAGACCATATAAGAAAAACAATACAAAATCAAGCGAAATCTGTTGATGTG  546

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACCCTCAGACCAGGTCTCAAA  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACCCTCAGACCAGGTCTCAAA  606

Qy   541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA  666

Qy   601 GCTCAAGCACAAAGCATTTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667 GCTCAAGCACAAAGCATTTTAAACAAAACCATCCAGGCTATACGATTTATGAACGTGAC  726

Qy   661 TCCTCAATCGTCACTCATGACAATGACAATGACATTTTCCGTACGACGATTTACCAATGGATCAAGAG  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727 TCCTCAATCGTCACTCATGACAATGACAATGACATTTTCCGTACGACGATTTACCAATGGATCAAGAG  786

Qy   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   787 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  846

Qy   781 AATGAGAAATAAACACACTGACCTGATCTCTGAGAAAATATTACGTCCTTAAAAAAGGG  840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   847 AATGAGAAATAAACACACTGACCTGATCTCTGAGAAAATATTACGTCCTTAAAAAAGGG  906

Qy   841 GAAAAGCCGTGATGATCCCTTTGATCGCAGTCACTGAAACTGTCACCATCAAATAACGTT  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   907 GAAAAGCCGTGATGATCCCTTTGATCGCAGTCACTGAAACTGTCACCATCAAATAACGTT  966

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   967 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1026

Qy   961 TTAGACTTCAGAGATTTATACGATCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1027 TTAGACTTCAGAGATTTATACGATCTCGTGATAAGGCTAAACTACTCTACACACAATCTC 1086

Qy  1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGGAGGATAATCACGATGAC 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1087 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGGAAGATAATCACGATGAC 1146

Qy  1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACACCGAAGGAGGAGAATGCTAGCTAT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1147 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1206

Qy  1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1207 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1266

Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA 1242
         ||||||||||||||||||||||||||||||||||||||||||
Db  1267 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA 1308

Search completed: May  6, 2004, 06:20:21
Job time : 5093 secs
```

OM nucleic - nucleic search, using sw model

```
              GenCore version 5.1.6
         Copyright  (c) 1993 - 2004 Compugen Ltd.
```

Run on: May   6, 2004, 03:54:48 ; Search time 4627 Seconds
                                          (without alignments)
                                  9467.629 Million cell updates/sec

Title: US-09-882-509-3
Perfect score: 1245
Sequence: 1 attgctggacctgagtggct.........ataaccctaacgacacaaataa 1245

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 1759305951B residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Pending_Patents_NA_Main:*
```
  1: /cgn2_6/ptodata72/pna/PCTUS_COMB.seq:*
  2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
  3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
  4: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*
  5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
  6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
  7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
  8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
  9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
 10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
 11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
 12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
 13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
 14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
 15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
 16: /cgn2_6/ptodata/2/pna/US092A_COMB.seq:*
 17: /cgn2_6/ptodata/2/pna/US092B_COMB.seq:*
 18: /cgn2_6/ptodata/2/pna/US093A_COMB.seq:*
 19: /cgn2_6/ptodata/2/pna/US093B_COMB.seq:*
 20: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
 21: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
 22: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
 23: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
 24: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
 25: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
 26: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
 27: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
 28: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
 29: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
 30: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
 31: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
 32: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
 33: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:*
 34: /cgn2_6/ptodata/2/pna/US098B_COMB.seq:*
 35: /cgn2_6/ptodata/2/pna/US098C_COMB.seq:*
 36: /cgn2_6/ptodata/2/pna/US098D_COMB.seq:*
 37: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:*
 38: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:*
 39: /cgn2_6/ptodata/2/pna/US099C_COMB.seq:*
 40: /cgn2_6/ptodata/2/pna/US099D_COMB.seq:*
 41: /cgn2_6/ptodata/2/pna/US099E_COMB.seq:*
 42: /cgn2_6/ptodata/2/pna/US099F_COMB.seq:*
 43: /cgn2_6/ptodata/2/pna/US100A_COMB.seq:*
 44: /cgn2_6/ptodata/2/pna/US100B_COMB.seq:*
 45: /cgn2_6/ptodata/2/pna/US101A_COMB.seq:*
 46: /cgn2_6/ptodata/2/pna/US101B_COMB.seq:*
 47: /cgn2_6/ptodata/2/pna/US102A_COMB.seq:*
 48: /cgn2_6/ptodata/2/pna/US102B_COMB.seq:*
 49: /cgn2_6/ptodata/2/pna/US103A_COMB.seq:*
 50: /cgn2_6/ptodata/2/pna/US103B_COMB.seq:*
 51: /cgn2_6/ptodata/2/pna/US104A_COMB.seq:*
 52: /cgn2_6/ptodata/2/pna/US104B_COMB.seq:*
 53: /cgn2_6/ptodata/2/pna/US106_COMB.seq:*
 54: /cgn2_6/ptodata/2/pna/US107A_COMB.seq:*
 55: /cgn2_6/ptodata/2/pna/US107B_COMB.seq:*
 56: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
 57: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
 58: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
 59: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
 60: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
 61: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
 62: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
 63: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
 64: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
 65: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
 66: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
 67: /cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
 68: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
 69: /cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
 70: /cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
 71: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
 72: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
 73: /cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
 74: /cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
 75: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
 76: /cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
 77: /cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
 78: /cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
 79: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:*
 80: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:*
 81: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
 82: /cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
 83: /cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
 84: /cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
 85: /cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
 86: /cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
 87: /cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
 88: /cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
 89: /cgn2_6/ptodata/2/pna/US6032_COMB.seq:*
 90: /cgn2_6/ptodata/2/pna/US6033_COMB.seq:*
 91: /cgn2_6/ptodata/2/pna/US6034_COMB.seq:*
 92: /cgn2_6/ptodata/2/pna/US6035_COMB.seq:*
 93: /cgn2_6/ptodata/2/pna/US6036_COMB.seq:*
 94: /cgn2_6/ptodata/2/pna/US6037_COMB.seq:*
 95: /cgn2_6/ptodata/2/pna/US6038_COMB.seq:*
 96: /cgn2_6/ptodata/2/pna/US6039_COMB.seq:*
 97: /cgn2_6/ptodata/2/pna/US6040_COMB.seq:*
 98: /cgn2_6/ptodata/2/pna/US6041_COMB.seq:*
 99: /cgn2_6/ptodata/2/pna/US6042_COMB.seq:*
100: /cgn2_6/ptodata/2/pna/US6043_COMB.seq:*
101: /cgn2_6/ptodata/2/pna/US6044_COMB.seq:*
102: /cgn2_6/ptodata/2/pna/US6045_COMB.seq:*
103: /cgn2_6/ptodata/2/pna/US6046_COMB.seq:*
104: /cgn2_6/ptodata/2/pna/US6047_COMB.seq:*
105: /cgn2_6/ptodata/2/pna/US6048_COMB.seq:*
106: /cgn2_6/ptodata/2/pna/US6049_COMB.seq:*
107: /cgn2_6/ptodata/2/pna/US6050_COMB.seq:*
108: /cgn2_6/ptodata/2/pna/US6051_COMB.seq:*
109: /cgn2_6/ptodata/2/pna/US6052_COMB.seq:*
110: /cgn2_6/ptodata/2/pna/US6053_COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1245	100.0	1245	5	US-08-145-684-1	Sequence 1, Appli
2	1245	100.0	1245	36	US-09-882-509-3	Sequence 1, Appli
3	1224.2	98.3	1245	20	US-09-471-349-1	Sequence 1, Appli
4	1224.2	98.3	1245	39	US-09-940-235-1	Sequence 1, Appli
5	1224.2	98.3	1245	53	US-10-631-558-1	Sequence 1, Appli
6	1222.6	98.2	1377	20	US-09-471-349-5	Sequence 5, Appli
7	1222.6	98.2	1377	39	US-09-940-235-5	Sequence 5, Appli
8	1222.6	98.2	1377	53	US-10-631-558-5	Sequence 5, Appli
9	1222.6	98.2	1782	20	US-09-471-349-11	Sequence 11, Appl
10	1222.6	98.2	1782	39	US-09-940-235-11	Sequence 11, Appl
11	1222.6	98.2	1782	53	US-10-631-558-11	Sequence 11, Appl
12	1221.2	98.1	1254	24	US-09-569-920-2	Sequence 2, Appli
13	1221.2	98.1	8893	24	US-09-569-920-3	Sequence 3, Appli
14	1209.8	97.2	1327	20	US-09-471-349-6	Sequence 6, Appli
15	1209.8	97.2	1327	39	US-09-940-235-6	Sequence 6, Appli
16	1209.8	97.2	1327	53	US-10-631-558-6	Sequence 6, Appli
17	1206.4	96.9	1209	27	US-09-658-681-11	Sequence 11, Appl
18	1203.8	96.7	1245	27	US-09-658-681-12	Sequence 12, Appl
19	1127.8	90.6	1661	20	US-09-471-349-10	Sequence 10, Appl
20	1127.8	90.6	1661	39	US-09-940-235-10	Sequence 10, Appl
21	1127.8	90.6	1661	53	US-10-631-558-10	Sequence 10, Appl
22	1124.6	90.3	2096	20	US-09-471-349-12	Sequence 12, Appl
23	1124.6	90.3	2096	39	US-09-940-235-12	Sequence 12, Appl
24	1124.6	90.3	2096	53	US-10-631-558-12	Sequence 12, Appl
25	1116	89.6	1122	27	US-09-658-681-13	Sequence 13, Appl
26	1116	89.6	1158	27	US-09-658-681-14	Sequence 14, Appl
27	1111.8	89.3	1541	20	US-09-471-349-9	Sequence 9, Appli
28	1111.8	89.3	1541	39	US-09-940-235-9	Sequence 9, Appli
29	1111.8	89.3	1541	53	US-10-631-558-9	Sequence 9, Appli
30	1093.2	87.8	1320	51	US-10-415-182A-8297	Sequence 8297, Ap
c 31	46.2	3.7	601	39	US-09-948-933-5546	Sequence 5546, Ap
c 32	46.2	3.7	2000	36	US-09-887-272A-5263	Sequence 5263, Ap
c 33	46.2	3.7	94499	82	US-60-258-275-180	Sequence 180, App
c 34	46.2	3.7	95139	39	US-09-948-933-606	Sequence 606, App
35	46.2	3.7	8059021	14	US-09-947-914-53	Sequence 53, Appl
c 36	44.4	3.6	98272	103	US-60-466-412-83647	Sequence 83647, A
c 37	44	3.5	1141	33	US-09-806-708A-22	Sequence 22, Appl
c 38	44	3.5	1141	33	US-09-806-708B-22	Sequence 22, Appl
39	43.8	3.5	2000	36	US-09-887-272A-5263	Sequence 5263, Ap
40	43.8	3.5	301828	39	US-09-949-016-13969	Sequence 13969, A
41	43.8	3.5	763446	104	US-09-470-166-8943	Sequence 8943, Ap
c 42	43.4	3.5	556	50	US-1-369-493-29120	Sequence 29120, A
43	43.4	3.5	556	93	US-60-360-039-29120	Sequence 29120, A
c 44	43.4	3.5	832	53	US-10-664-025-2813	Sequence 2813, Ap
45	43.4	3.5	832	70	US-60-147-499-2813	Sequence 2813, Ap

ALIGNMENTS

RESULT 1
US-08-145-684-1
; Sequence 1, Application US/08145684
; GENERAL INFORMATION:
; APPLICANT: Garcia, M.P.E. et al
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND EXPRESSION
; TITLE OF INVENTION: OF A GENE WHICH CODES FOR STREPTOKINASE, NUCLEOTIDE SEQUENCE
; TITLE OF INVENTION: OBTAINED, RECOMBINANT DNA AND TRANSFORMED MICROORGANISMS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stanger, Michaelson, Spivak and Wallace, Esq.
; STREET: Parkway 109 Office Center, 328 Newman Springs Road,
; STREET: P. O. Box 8489
; CITY: Red Bank
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2" 1.44Mb IBM compatible diskette

COMPUTER: IBM PS/2 Model 80
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: Microsoft Word for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/145,684
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/703,778
FILING DATE: May 22, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Michaelson, Peter L.
REGISTRATION NUMBER: 30090
REFERENCE/DOCKET NUMBER: Centro-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)530-6671
TELEFAX: (908)530-6584
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1245 base pairs
TYPE: Nucleotide with corresponding protein
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Streptococcus equisimilis from group C of Lanfield
ORGANISM: definition
IMMEDIATE SOURCE: ATCC-9542 strain
FEATURE: from 1 to 1245 bp mature peptide
OTHER INFORMATION:
OTHER INFORMATION: Properties: Streptokinase gene
OTHER INFORMATION: The gene product binds to human plasminogen
OTHER INFORMATION: The gene product is an activator of human plasminogen

US-08-145-684-1

Query Match   100.0%; Score 1245; DB 5; Length 1245;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1245; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 ATTGCTGGACCTGGAGTGGCTGCTAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATTGCTGGACCTGGAGTGGCTGCTAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT   60

Qy   61 AGCGTTGCTGCTGGTACTGTTGAGGGGACGAACATCAAGACATTAGTCTTAAATTTTTGAAATT  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGCGTTGCTGCTGGTACTGTTGAGGGGACGAACATCAAGACATTAGTCTTAAATTTTTGAAATT  120

Qy  121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGAGCAAGGCTTAAGTCCAAAATCA  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGAGCAAGGCTTAAGTCCAAAATCA  180

Qy  181 AAACCATTTGCTACTGATAGTGGCGCGGATCGCTAACGTCCACAGTAACGACGACTTTGAGGTC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AAACCATTTGCTACTGATAGTGGCGCGGATCGCTAACGTCCACAGTAACGACGACTTTGAGGTC  240

Qy  241 AAGGCTATTCAAGAACAATTGATCGTCAACAATTACTGATCGAAACGGCAAGGTCTACTTGCTGAC  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AAGGCTATTCAAGAACAATTGATCGTCAACAATTACTGATCGAAACGGCAAGGTCTACTTGCTGAC  300

Qy  301 ATTGATTTTGCAAGCGATGCAACCATTCAGCAACCATTACTGTGAAACGGAAACGGCAAGGGCTACTTTGCTGAC  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ATTGATTTTGCAAGCGATGCAACCATTCAGCAACCATTACTGTGAAACGGAAACGGCAAGGGCTACTTTGCTGAC  360

Qy  361 AAAGATGTTCGGTAACCTTGCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AAAGATGTTCGGTAACCTTGCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT  420

Qy  421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG  480

Qy  481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACACCAGTCTCAAA  540
```

Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1245;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Db   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACACCAGGTCTCAAA 540

Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTAGTTGTT   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTAGTTGTT   60

Qy    61 AGCGTTGCTGGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AGCGTTGCTGGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT  120

Qy   121 GACCTAAACATCACGACCTGCTCATGGAGGAAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GACCTAAACATCACGACCTGCTCATGGAGGAAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180

Qy   181 AAACCATTTGCTACTGATAGTGCGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAACCATTTGCTACTGATAGTGCGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA  240

Qy   241 AAGGCTATTCAAGAACAATGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AAGGCTATTCAAGAACAATGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGACTGAAACGGCAAGGTCTACTTTGCTGAC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATTGATTTTGCAAGCGATGCAACCATTACTGACTGAAACGGCAAGGTCTACTTTGCTGAC  360

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420

Qy   421 GTGGCGGTTAGACCATATAAAGAGAAAAAAAACCAATACAAAATCAAGCCGAAATCTGTTGATGTG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTGGCGGTTAGACCATATAAAGAGAAAAAAAACCAATACAAAATCAAGCCGAAATCTGTTGATGTG  480

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  540

Qy   541 GATACTAAGCTATTGAAAACACTAGCTATCGTTGGTGACACCATCACATCTCAAGAATTACTA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GATACTAAGCTATTGAAAACACTAGCTATCGTTGGTGACACCATCACATCTCAAGAATTACTA  600

Qy   601 GCTCAAGCACAAAAGCATTTAAACAAAATGCATTTTCGTACGACGATTTTACCAATGAACGTGAC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCTCAAGCACAAAAGCATTTAAACAAAATGCATTTTCGTACGACGATTTTACCAATGAACGTGAC  660

Qy   661 TCCTCAATCGTCACTCATGACAATGAACATTTTCCGTACGACTATTACCAATGGAATGCTAGCTAT  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCCTCAATCGTCACTCATGACAATGAACATTTTCCGTACGACTATTACCAATGGAATGCTAGCTAT  720

Qy   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGGATCAATCAAAAAATCTGGTCTG  780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGGATCAATCAAAAAATCTGGTCTG  780

Qy   781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAAGGG  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAAGGG  840

Qy   841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACGTGTTCACCATCAAATACGTT  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACGTGTTCACCATCAAATACGTT  900

Qy   901 GATGTCAACACCAACGAATTGCTAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GATGTCAACACCAACGAATTGCTAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC  960

Qy   961 TTAGACTTCAGAGATTTATGGTTATATACGATCCTGTGATAAGGCTAAACTACTTCACAACAATCTC 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTAGACTTCAGAGATTTATGGTTATATACGATCCTGTGATAAGGCTAAACTACTTCACAACAATCTC 1020

Qy  1021 GATGCGTTTTGGTATTATGGACTATACCTTAACTGACAGAATAATCACGATGAC 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 GATGCGTTTTGGTATTATGGACTATACCTTAACTGACAGAATAATCACGATGAC 1080

Qy  1081 ACCAACCGTATCATCATAACCGTTTATATGGGCAAGCGACCCGAAGCGAGAGAGATGCTAGCTAT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACCAACCGTATCATCATAACCGTTTATATGGGCAAGCGACCCGAAGCGAGAGAGATGCTAGCTAT 1140

Qy  1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAGAGTTTACAGCTACCTG 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAGAGTTTACAGCTACCTG 1200

Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
           |||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
```

RESULT 2
US-09-882-509-3
; Sequence 3, Application US/09882509
; GENERAL INFORMATION:
; APPLICANT: Kuppusamy, Mosuvan
; APPLICANT: Srinivas, Vellimedu K
; APPLICANT: Lahiri, Subhra
; APPLICANT: Ella, Krishna
; APPLICANT: Khatri, Ghan S
; TITLE OF INVENTION: Recombinant Streptokinase
; FILE REFERENCE: 51321.003
; CURRENT APPLICATION NUMBER: US/09/882,509
; CURRENT FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis (ATCC 9542)
US-09-882-509-3

Query Match   100.0%;   Score 1245;   DB 36;   Length 1245;

```
Qy  361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420

Qy  421 GTGCGCGTTAGACCATATAAAGAGAAAAACCAATACAACAAAATCAAGCGAAATCTGTTGATGTG  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTGCGCGTTAGACCATATAAAGAGAAAAACCAATACAACAAAACCAAGCGAAATCTGTTGATGTG  480

Qy  481 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GAATATACTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  540

Qy  541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACCATCACATCTCAAGAATTACTA  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACCATCACATCTCAAGAATTACTA  600

Qy  601 GCTCAAGCACAAAGCATTTAAACAAACCCACCCAGCTATACGATTTATGAACGTGAC  660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC  660

Qy  661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG  720

Qy  721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TTTACTTACCGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG  780

Qy  781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATTATTACGTCCTAAAAAAGGG  840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATTATTACGTCCTAAAAAAGGG  840

Qy  841 GAAAAGCCGTATGATCCCTTGATGCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GAAAAGCCGTATGATCCCTTTGATGCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT  900

Qy  901 GATGTCAACACCAAGAATTGCTAAAAAGGAGCAGCTCTCTAACAGCTAGCGAACGTAAC  960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 GATGTCGATACCAACGAATTGCTAAAAAGTGAACAGCGTCTTAACAGCTAGCGAACGTAAC  960

Qy  961 TTAGACTTCAGAGATTTATACGATCCTCGTGATACGATCCTAACTACTCTTACAACAATCTC  1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTAGACTTCAGAGATTTATACGATCCTCGTGATACGATCCTAACTACTCTTACAACAATCTC  1020

Qy 1021 GATGCTTTGGTATTATGGACTATACCTTAACTGAAAAGTAGAGGATAATCACGATGAC  1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 GATGCTTTGGTATTATGGACTATACCTTAACTGAAAAGTAGAGGATAATCACGATGAC  1080

Qy 1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT  1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT  1140

Qy 1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGACAGAGAACCAGAAGTTTACAGCTACCTG  1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGACAGAGAACCAGAAGTTTACAGCTACCTG  1200

Qy 1201 CGTTATACAGGGACACCTATACCCTAACGACAAATAA  1245
        |||||||||||||||||||||||||||||||||||||
Db 1201 CGTTATACAGGGACACCTATACCCTAACGACAAATAA  1245
```

RESULT 3
US-09-471-349-1
US-09-471-349-1

Query Match   98.3%; Score 1224.2; DB 20; Length 1245;
Best Local Similarity   99.0%; Pred. No. 0;
Matches 1232; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy    1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT  60

Qy   61 AGCGGTTGCTGTGGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTGAAATT  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGCGGTTGCTGTGGTACTGTTGAGGGGACGAATCAAGACATTAGTTCTTAAATTTTTGAAATC  120

Qy  121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180

Qy  181 AAACCATTGCTACTGATAGTGGCGGCGATGCCACATAAAACTTGAAAAAGTGACTTACTA  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AAACCATTGCTACTGATAGTGGCGGCGATGCCACATAAAACTTGAGAAAGCTGACTTACTA  240

Qy  241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300

Qy  301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAATGCTCTACTTTGCTGAC  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GATCGTTGCTGGTAAGCGATGCAACCATTACTGATCGAAATGCTCTACTTTGCTGAC  360
```

RESULT 4
US-09-940-235-1

```
;  TITLE OF INVENTION:   PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
;  TITLE OF INVENTION:   CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
;  TITLE OF INVENTION:   PROTEIN
;  FILE REFERENCE: 07064-009002
;  CURRENT APPLICATION NUMBER: US/09/940,235
;  CURRENT FILING DATE: 2002-04-09
;  PRIOR APPLICATION NUMBER: 09/471,349
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: IN 3825/DEL/98
;  PRIOR FILING DATE: 1998-12-24
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 1
;  LENGTH: 1245
;  TYPE: DNA
;  ORGANISM: Streptococcus equisimilis
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)...(1242)
US-09-940-235-1

Query Match             98.3%;  Score 1224.2;  DB 39;  Length 1245;
Best Local Similarity   99.0%;  Pred. No. 0;
Matches  1232;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT 60

Qy     61 AGCGGTTGCTGGTACTGTTGAGGGGACGAATCAACATTAGTCTTAAATTTTTTGAAATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGCGGTTGCTGGTACTGTTGAGGGGACGAATCAACATTAGTCTTAAATTTTTTGAAATC 120

Qy    121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180

Qy    181 AAACCATTTGCTACTGATAGTGGCGGCGATGGCCATCATAAACTTGAAAAAGCTGACTTACTA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAACCATTTGCTACTGATAGTGGCGGCGATGGCCATCATAAACTTGAAAAAGCTGACTTACTA 240

Qy    241 AAGGCTATTCAAGAACAATGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AAGGCTATTCAAGAACAATGATCGCTAACGTCACAGTAACGACGACTACTTTGAGGGTC 300

Qy    301 ATTGATTTTGCAAGCGGATGCAACCATTACTGATCGAAACCGCAAGGTCTTCACTTTGCTGAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATTGATTTTGCAAGCGGATGCAACCATTACTGATCGAAACCGCAAGGTCTTCACTTTGCTGAC 360

Qy    361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTTCCAAGAATTTTGCTAAGCGGGACAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTTCCAAGAATTTTGCTAAGCGGGACAT 420

Qy    421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480

Qy    481 GAATATACTGTACAGTTTACTCCCCTTAAACCCTGATGACCAGGTCTCAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAATATACTGTACAGTTTACTCCCCTTAAACCCTGATGACCAGGTCTCAAA 540

Qy    541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACCATCACATCTCAAGAATTACTA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACCATCACATCTCAAGAATTACTA 600

Qy    601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCACCGGCTATACGATTTATGAACGTGAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCACCGGCTATACGATTTATGAACGTGAC 660

Qy    661 TCCTCAATCGTCACTCATGACAATGACAATTTCCGTACGATTTTACCAATGGATCAAGAG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCCTCAATCGTCACTCATGACAATGACAATTTCCGTACGATTTTACCAATGGATCAAGAG 720

Qy    721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780

Qy    781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840

Qy    841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900

Qy    901 GATGTCAACACCAACGAATTGCTAAAAAGCGCAGCGCTCTTAACAGCTAGCGAACGTAAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GATGTCGATACCGATACCGAATTGCTAAAAAGTGCTAAAAAGTGAGCGCTCTTAACAGCTAGCGAACGTAAC 960

Qy    961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTCTTACAACAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTCTTACAACAATCTC 1020

Qy   1021 GATGCTTTTGGTATTATGGACTATATCCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GATGCTTTTGGTATTATGGACTATATCCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1080

Qy   1081 ACCAACCGTATCATAACCGTTTATGTGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACCAACCGTATCATAACCGTTTATGTGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1140

Qy   1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTACAGCTACCTG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTACAGCTACCTG 1200

Qy   1201 CGTTATACAGGGACGACCTATACCTGATAACCCTAACGACACAAATAA 1245
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CGTTATACAGGGACGACCTATACCTGATAACCCTAACGACACAAATAA 1245


RESULT 5
US-10-631-558-1
;  Sequence 1, Application US/10631558
;  GENERAL INFORMATION:
;  APPLICANT: Kumar, Rajesh
;  APPLICANT: Sahni, Girish
;  APPLICANT: Roy, Chait
;  APPLICANT: Rajagopal, Kammara
;  APPLICANT: Nihalani, Deepak
;  APPLICANT: Sundaram, Vasudha
;  APPLICANT: Yadav, Mahavir
;  TITLE OF INVENTION:   NOVEL CLOT-SPECIFIC STREPTOKINASE
;  TITLE OF INVENTION:   PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
;  TITLE OF INVENTION:   CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
;  TITLE OF INVENTION:   PROTEIN
;  FILE REFERENCE: 07064-009002
;  CURRENT APPLICATION NUMBER: US/10/631,558
;  CURRENT FILING DATE: 2003-07-31
;  PRIOR APPLICATION NUMBER: US/09/940,235
;  PRIOR FILING DATE: 2002-04-09
;  PRIOR APPLICATION NUMBER: 09/471,349
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: IN 3825/DEL/98
;  PRIOR FILING DATE: 1998-12-24
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 1
;  LENGTH: 1245
;  TYPE: DNA
;  ORGANISM: Streptococcus equisimilis
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)...(1242)
US-10-631-558-1

Query Match             98.3%;  Score 1224.2;  DB 53;  Length 1245;
```

Best Local Similarity 99.0%;   Pred. No. 0;
Matches 1232; Conservative   0; Mismatches 13; Indels 0; Gaps 0;

```
Qy    1  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   60

Qy   61  AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATT  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATC  120

Qy  121  GACCTAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GATCTAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180

Qy  181  AAACCATTTGCTACTGATAGTGCGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  AAACCATTTGCTACTGATAGTGCGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA  240

Qy  241  AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300

Qy  301  ATTGATTTTGCAAGCGATGCAACCATTACTGACTCGGTCTG  360
         |||||||||||||||||||||||||||||||||||||||||
Db  301  ATTGATTTTGCAAGCGATGCAACCATTACTGACTCGGTCTG  360

Qy  361  AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420

Qy  421  GTGCGCGGTTAGACCATATAAAGAAAATCAAGCGAAATCTGTTGATGTG  480
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GTGCGCGGTTAGACCATATAAAGAAAATCAAGCGAAATCTGTTGATGTG  480

Qy  481  GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGAGTCTCAAA  540
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGAGTCTCAAA  540

Qy  541  GATACTAAGCTATTGAAAACACTAGCTATCGGTGCACACCATCACATCTCAAGAATTACTA  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  GATACTAAGCTATTGAAAACACTAGCTATCGGTGCACACCATCACATCTCAAGAATTACTA  600

Qy  601  GCTCAAGCACAAAAGCATTTTAAACAAAAACAACCCAGGCTATACGATTTATGAACGTGAC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  GCTCAAGCACAAAAGCATTTTAAACAAAAACAACCCAGGCTATACGATTTATGAACGTGAC  660

Qy  661  TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGATCAAGAG  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGATCAAGAG  720

Qy  721  TTTACTTACCATGTCAAAATGCTAAAAGCTAAAAAATCGGAACAAGCTTATAGGATCAAT  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  TTTACTTACCGTGTTAAAAATCGGAACAAGCTTATAGGATCAAT  780

Qy  781  AATGAAGAAATAAACAACAATCGGAGAAATATTACGGTCCCTTAAAAAGGG  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  AATGAAGAAATAAACAACAATCGGAGAAATATTACGGTCCCTTAAAAAGGG  840

Qy  841  GAAAAGCCGTATGATCCCTTTGATCGCAGTTGTCACCATCAAATACGTT  900
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  GAAAAGCCGTATGATCCCTTTGATCGCAGTTGTCACCATCAAATACGTT  900

Qy  901  GATGTCAACACCACCAAGCGATTGCTAAAAAGCAGCTGCTAAACAGCTAGCAGCGAACGTAAC  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  GATGTCGATACCACCAAGCGATTGCTAAAAAGCAGCTCTTAACAGCTAGCAGCGAACGTAAC  960

Qy  961  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020

Qy 1021  GATGCTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  GATGCTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1080

Qy 1081  ACCAACCGTATCATAACCGTTTATATGGCCAAGCGACCCGAGGAGGAGAATGCTAGCTAT 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  ACCAACCGTATCATAACCGTTTATATGGCCAAGCGACCCGAGGAGGAGAATGCTAGCTAT 1140

Qy 1141  CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200

Qy 1201  CGTTATACAGGGACGACACCTATACCTGATGAATAACCCTAACGACAAATAA 1245
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201  CGTTATACAGGGACGACACCTATACCTGATGAATAACCCTAACGACAAATAA 1245
```

RESULT 6
US-09-471-349-5
; Sequence 5, Application US/09471349
; GENERAL INFORMATION:
; APPLICANT: Sahni, Girish
; APPLICANT: Kumar, Rajesh
; APPLICANT: Roy, Chaiti
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE PROTEINS POSSESSING ALTERED
; TITLE OF INVENTION: PLASMINOGEN ACTIVATION CHARACTERISTICS AND A PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF SAID PROTEINS
; FILE REFERENCE: 07064/009001
; CURRENT APPLICATION NUMBER: US/09/471,349
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-471-349-5

Query Match   98.2%; Score 1222.6; DB 20; Length 1377;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
Qy    1  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  133  ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTGGTTGTT  192

Qy   61  AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATT  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  193  AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATC  252

Qy  121  GACCTAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  253  GATCTAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA  312

Qy  181  AAACCATTTGCTACTGATAGTGCGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  313  AAACCATTTGCTACTGATAGTGCGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA  372

Qy  241  AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  373  AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC  432

Qy  301  ATTGATTTTGCAAGCGATGCAACCATTACTGACTCGGTCTGAC  360
         |||||||||||||||||||||||||||||||||||||||||||
Db  433  ATTGATTTTGCAAGCGATGCAACCATTACTGACTCGGTCTGAC  492

Qy  361  AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  493  AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT  552
```

FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/09/940,235
CURRENT FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 5
  LENGTH: 1377
  TYPE: DNA
  ORGANISM: Streptococcus equisimilis
US-09-940-235-5

Query Match   98.2%; Score 1222.6; DB 39; Length 1377;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   133 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTGGTTGTT  192

Qy    61 AGCGGTTGCTGGTACGTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   193 AGCGGTTGCTGGTACGTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC  252

Qy   121 GACTTAACATCACGACCTGCTCATGAGGAAAGACAGAGAGCAAGGCTTAAGTCCAAAATCA  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253 GATCTTAACATCACGACCTGCTCATGAGGAAAGACAGAGAGCAAGGCTTAAGTCCAAAATCA  312

Qy   181 AAACCATTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTACTA      240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   313 AAACCATTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTACTA      372

Qy   241 AAGGCTATTCAAGAACAATTGATCGTCAAGTCGTAACGTCCACAGTAACGACCACTACTTGAGGTC  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   373 AAGGCTATTCAAGAACAATTGATCGTCAAGTCGTAACGTCCACAGTAACGACCACTACTTGAGGTC  432

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGATAATCACGCAAGTCTCTACTTGCTGAC  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   433 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGATAATCACGCAAGTCTCTACTTGCTGAC  492

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTTCCAAGAATTTTGCTAAGCCGGACAT  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   493 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTTCCAAGAATTTTGCTAAGCCGGACAT  552

Qy   421 GTGCGCGGTTAGACCATATAAAGAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG  480
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   553 GTGCGCGGTTAGACCATATAAAGAGAAAAACCAATACAAAACCAAGCGAAATCTGTTGATGTG  612

Qy   481 GAATATACTGTACAGTTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   613 GAATATACTGTACAGTTGTACAGTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA  672

Qy   541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   673 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA  732

Qy   601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC     660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   733 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC     792

Qy   661 TCCTCAATGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG   720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   793 TCCTCAATGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG   852

Qy   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  780
           ||||||||||  || |||||||||||||||||||| |  |||||||||||||  |||||
Db   853 TTTACTTACCGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCGGTCTG   912

Qy   781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   913 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG  972

Qy   841 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTTGACTTCACCTCACAAATACGTT     900
           ||||||||||||||||||||||||||||||| |||||||||||| |||||||||||
Db   973 GAAAAGCCGTATGATCCCTTGATCGCAGTCAGTTGACTTCACCATCAAATACGTT    1032

Qy   901 GATGTCAACACCAACGAATTGCTAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAAC   960
           ||||||  |||||||||||||||||||| |||||||||||||||||||||||||||||
Db  1033 GATGTCGAATACCAACGAATTGCTAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC  1092

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATGATAAGGCTCAAACTACTCTACAACATCTC 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db  1093 TTAGACTTCAGAGATTTATACGATCCTCGTGATGATAAGGCTCAAACTACTCTACAACAATCTC 1152

Qy  1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGAAAAGTAGAGGATAATCACGATGAC  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1153 GATGCTTTTGGTATTATGGACTATACCTTAACTGAAAAGTAGAGGATAATCACGATGAC  1212

Qy  1081 ACCAACCGTATCATAACCGTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1213 ACCAACCGTATCATAACCGTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1272

Qy  1141 CATTTAGCCTATGATAAAGATCGTTATACCGGAAGAACGAGAGAGTTTACAGCTGACTACCTG 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1273 CATTTAGCCTATGATAAAGATCGTTATACCGGAAGAACGAGAGAGTTTACAGCTGACTACCTG 1332

Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
           |||||||||||||||||||||||||||||||||||||||||||||
Db  1333 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1377
```

RESULT 7
US-09-940-235-5
; Sequence 5, Application US/09940235
; GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammara
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN

```
Db  913 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 972      Db  193 AGGCTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC 252
                                                                               Qy  121 GACCTAACATCACGACCTGCTCACTTGAAACTGTTCACCATCAAATACGTT 180
Qy  841 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||||                 Db  253 GATCTAACATCACGACCTGCTCATGGCGCGGATGGCGCGGATGGCGCGGATGGCTAACTA 312
Db  973 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 1032
                                                                               Qy  181 AAACCATTTGCTACTGATAGTGGCGCGGATGGCGCGGATGGCTAACTA 240
Qy  901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGGGAACGTAAC 960         ||||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||                 Db  313 AAACCATTTGCTACTGATAGTGGCGCGGATGGCGCGGATGGCTAACTA 372
Db  1033 GATGTCGATACCAACGAATTGCTAAAAAGTAGAGGATCAAGCTAGGGAACGTAAC 1092
                                                                               Qy  241 AAGGCTATTCAAGAACAATTCAAGAACAATTCTTGAGGTC 300
Qy  961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020       ||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||                 Db  373 AAGGCTATTCAAGAACAATTCAAGAACAATTCTTGAGGTC 432
Db  1093 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1152
                                                                               Qy  301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGCGAGGTCTACTTTGCTGAC 360
Qy  1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGACAAGGAGGATAATCACGATGAC 1080         ||||||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||                 Db  433 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGCGAGGTCTACTTTGCTGAC 492
Db  1153 GATGCTTTTGGTATTATGGACTATACCTTAACTGACAAGGAGGATAATCACGATGAC 1212
                                                                               Qy  361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420
Qy  1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1140       ||||||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||                 Db  493 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 552
Db  1213 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1272
                                                                               Qy  421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCGTTGATGTG 480
Qy  1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGCAGAAGTTTACAGCTACCTG 1200     ||||||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||                 Db  553 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCGTTGATGTG 612
Db  1273 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGCAGAAGTTTACAGCTACCTG 1332
                                                                               Qy  481 GAATATACTGTACAGTTTACTCCCTTAAACCCTAACCCTGATGACGATTTCAGACCAGGTCTCAAA 540
Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245                       |||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||                       Db  613 GAATATACTGTACAGTTTACTCCCTTAAACCCTAACCCTGATGACGATTTCAGACCAGGTCTCAAA 672
Db  1333 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1377
                                                                               Qy  541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 600
                                                                                       ||||||||||||||||||||||||||||||||||||||||||||||||||
                                                                               Db  673 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 732

                                                                               Qy  601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGTATTGAACGTGAC 660
RESULT 8                                                                               ||||||||||||||||||||||||||||||||||||||||||||||||||
US-10-631-558-5                                                                Db  733 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGTATTGAACGTGAC 792
; Sequence 5, Application US/10631558
; GENERAL INFORMATION:                                                         Qy  661 TCCTCAATCGTCACTCATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 720
; APPLICANT: Kumar, Rajesh                                                             ||||||||||||||||||||||||||||||||||||||||||||||||||
; APPLICANT: Sahni, Girish                                                     Db  793 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 852
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara                                                Qy  721 TTTACTTACCATGTCAAAAATCGGAACAAGCTTATGATGATCAATAAAAAATCGGTCTG 780
; APPLICANT: Nihalani, Deepak                                                          ||||||||||||||||||||||||||||||||||||||||||||||||||
; APPLICANT: Sundaram, Vasudha                                                 Db  853 TTTACTTACCGTGTAAAAATCGGAACAAGCTTATAGGATCAATAAAAAATCGGTCTG 912
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE                         Qy  781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION                ||||||||||||||||||||||||||||||||||||||||||||||||||
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID Db  913 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 972
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002                                                 Qy  841 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900
; CURRENT APPLICATION NUMBER: US/10/631,558                                             ||||||||||||||||||||||||||||||||||||||||||||||||||
; CURRENT FILING DATE: 2003-07-31                                              Db  973 GAAAAGCCGTATGATCCCTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 1032
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09                                                Qy  901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGGGAACGTAAC 960
; PRIOR APPLICATION NUMBER: 09/471,349                                                 ||||||||||||||||||||||||||||||||||||||||||||||||||
; PRIOR FILING DATE: 1999-12-23                                                Db  1033 GATGTCGATACCAACGAATTGCTAAAAAGTAGAGGATCAAGCTAGGGAACGTAAC 1092
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24                                                Qy  961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
; NUMBER OF SEQ ID NOS: 28                                                              ||||||||||||||||||||||||||||||||||||||||||||||||||
; SOFTWARE: FastSEQ for Windows Version 4.0                                    Db  1093 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1152
; SEQ ID NO 5
; LENGTH: 1377                                                                 Qy  1021 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1080
; TYPE: DNA                                                                             ||||||||||||||||||||||||||||||||||||||||||||||||||
; ORGANISM: Streptococcus equisimilis                                         Db  1153 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1212
US-10-631-558-5
                                                                               Qy  1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1140
Query Match    98.2%;  Score 1222.6;  DB 53;  Length 1377;                             ||||||||||||||||||||||||||||||||||||||||||||||||||
Best Local Similarity  98.9%;  Pred. No. 0;                                    Db  1213 ACCAACCGTATCATAACCGTTCCATCTGTCAACAACAGCCAATTAGTTGTT 1272
Matches 1231;  Conservative  0;  Mismatches  14;  Indels  0;  Gaps  0;
                                                                               Qy  1 ATTGCTGGACCTGACGTGGCTGCTAGAACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
Qy  1 ATTGCTGGACCTGACGTGGCTGCTAGAACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60           ||||||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||                 Db  133 ATTGCTGGACCTGACGTGGCTGCTAGAACCGTCCATCTGTCAACAACAGCCAATTGGTTGTT 192
Db  133 ATTGCTGGACCTGACGTGGCTGCTAGAACCGTCCATCTGTCAACAACAGCCAATTGGTTGTT 192
                                                                               Qy  61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
Qy  61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120             ||||||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||                 Db  1273 AGGCTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 1332
```

```
Db  1018 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA 1077
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541 GATACTAAGCTAGTGAAAAACACTAGCTATGGTGACACCATCACATCTCAAGAATTACTA 600
         |||||||||||| ||||||||||||||||| |||||||||||||||||||||||||||||
Db  1078 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 1137

Qy   601 GCTCAAGCACAAAGCATTTTAAACAAAACCACCCAGCTATACGATTTATGAACGTGAC 660
         |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  1138 GCTCAAGCACAAAGCATTTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC 1197

Qy   661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1198 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG 1257

Qy   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780
         |||||||||| |||| ||||||||||||||||||| |||||||||||  ||||||||||
Db  1258 TTTACTTACCGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAAATCTGGTCTG 1317

Qy   781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCCTTAAAAAGGG 840
         |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  1318 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377

Qy   841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTGTTCACCATCAAATACGTT 900
         |||||||||||||||||||||||||||||||||  |||||||||||||||||
Db  1378 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAACTGTTCACCATCAAATACGTT 1437

Qy   901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGGAACGTAAC 960
         ||||| || |||||||||||||||||||| ||||||||||||||||||||| |||||||
Db  1438 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGGGAACGTAAC 1497

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1498 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1557

Qy  1021 GATGCTTTTGGTATTATGATGAATCATAACCGTTATATGGGACAAGCGACCCGAAGGACGA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1558 GATGCTTTTGGTATTATGATGAATCATAACCGTTATATGGGACAAGCGACCCGAAGGACGA 1617

Qy  1081 ACCAACCGTATCATAACCGTTATATGGGCAAGCGACCACCGGAAGGAGAATGCTAGCTAT 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1618 ACCAACCGTATCATAACCGTTATATGGGCAAGCGACCACCGGAAGGAGAATGCTAGCTAT 1677

Qy  1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGATCGTTATACCAGCTACCTG 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGATCGTTATACAGCTACCTG 1737

Qy  1201 CGTTATACAGGGACACCTATACCTGATAACCTAACGACAAATAA 1245
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1738 CGTTATACAGGGACACCTATACCTGATAACCTAACGACAAATAA 1782
```

RESULT 9

Query Match   98.2%;   Score 1222.6;   DB 20;   Length 1782;
Best Local Similarity   98.9%;   Pred. No. 0;
Matches 1231;   Conservative 0;   Mismatches 14;   Indels 0;   Gaps 0;

```
Qy     1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTTCCATCTGTCAACACAGCCAATTAGTTGTT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTTCCATCTGTCAACACAGCCAATTGGTTGTT 597

Qy    61 AGGGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
         || ||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db   598 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC 657

Qy   121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGGCAAGGCTTAAGTCCAAAATCA 180
         || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   658 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGGCAAGGCTTAAGTCCAAAATCA 717

Qy   181 AAACCATTTGCTACTGATAGTGGCGGCGATGCCACATAAACTTGAAAAAGCTGACTACTA 240
         ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   718 AAACCATTTGCTACTGATAGTGGCGGCGATGTCACATAAACTTGAAAAAGCTGACTACTA 777

Qy   241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   778 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 837

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTACTGATCGAAACGGCAAGGTCTACTTGCTGAC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   838 ATTGATTTTGCAAGCGATGCAACCATTACTACTGATCGAAACGGCAAGGTCTACTTGCTGAC 897

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   898 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT 957

Qy   421 GTGCGCGTTAGACCATATAAAGAAAAAACAATACAAAATCAAGCGAAATCTGTTGATGTG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   958 GTGCGCGTTAGACCATATAAAGAAAAAACAATACAAAATCAAGCGAAATCTGTTGATGTG 1017

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 10

```
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 11
;  LENGTH: 1782
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Hybrid cassette
US-09-940-235-11

Query Match      98.2%;     Score 1222.6;    DB 39;   Length 1782;
Best Local Similarity  98.9%;    Pred. No. 0;
Matches 1231;    Conservative   0;   Mismatches    14;   Indels     0;   Gaps     0;

Qy       1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAATTGGTTGTT 597

Qy      61 AGCGGTTGCTCGTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     598 AGCGGTTGCTCGTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAATC 657

Qy     121 GACCTAACATCACGACACCGTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     658 GATCTAACATCACGACACCGTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 717

Qy     181 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTGAAAAAGCTGACTTACTA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     718 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTGAAGAAAGCTGACTTACTA 777

Qy     241 AAGGCTATTCAAGAACAATTGATCGTAAGCTCCACAGTAACGAATCTGTTGAGGTC 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     778 AAGGCTATTCAAGAACAATTGATCGTAACGTCCACAGTAACGACTACTTTGAGGTC 837

Qy     301 ATTGATTTTGCAAGGCGATGCAACCATTACTGATCGAAACGGCCAAGGCTCTACTTTGCTGAC 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     838 ATTGATTTTGCAAGGCGATGCAACCATTACTGATCGAAACGGCCAAGGTCTCACTTTGCTGAC 897

Qy     361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCGTCCAAGAATTTTTGCTAAGCGGACAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     898 AAAGATGGTTCGGTAACCTTGCCGACCCAACCGTCCAAGAATTTTTGCTAAGCGGACAT 957

Qy     421 GTGCGCCGTTAGACCATATAAGAGAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     958 GTGCGCGGTTAGACCATATAAAGAGAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 1017

Qy     481 GAATATACTGTACAGTTTACTCCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1018 GAATATACTGTACAGTTTACTCCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA 1077

Qy     541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACCATCACATCTCAAGAATTACTA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1078 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACCATCACATCTCAAGAATTACTA 1137

Qy     601 GCTCAAGCACAAAAGCATTTAAACAAAACCCAGGCTATACGATTTATGAACGTGAC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1138 GCTCAAGCACACAAAGCATTTAAACAAAACCCAGGCTATACGATTTATGAACGTGAC 1197

Qy     661 TCCTCAATCGTCACTCGTCATGCATGACAATGACAATTTCCGTACGATTTACCAATCGATCAAGAG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1198 TCCTCAATCGTCACTCGTCATGCATGACAATGACAATTTCCGTACGATTTACCAATCGATCAAGAG 1257

Qy     721 TTTACTTACCGTGTGTTAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1258 TTTACTTACCGTGTGTTAAAAATCGGGAACAAGCTTATAGGAATAATAAAAAAATCTGGTCTG 1317

Qy     781 AATGAAGAAATAAACAACAACTGACCTGACCTCTGAGAAAATATTACGTTCCTTAAAAAAGGG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1318 AATGAAGAAATAAACAACAACTGACCTGACCTCTGAGAAAATATTACGTTCTTAAAAAAGGG 1377

Qy     841 GAAAAGGCTATGATCCCTTTGATCGGCAGTCACTTGAAACTGTTCACCATCAAAATACGTT 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1378 GAAAAGGCTATGATCCCTTTGATCGGCAGTCACTTGAAACTGTTCACCATCAAAATACGTTCA 1437
```

RESULT 11
US-10-631-558-11
; Sequence 11, Application US/10631558
; GENERAL INFORMATION:
;   APPLICANT: Kumar, Rajesh
;   APPLICANT: Sahni, Girish
;   APPLICANT: Roy, Chait
;   APPLICANT: Rajagopal, Kammara
;   APPLICANT: Nihalani, Deepak
;   APPLICANT: Sundaram, Vasudha
;   APPLICANT: Yadav, Mahavir
;   TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
;   TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
;   TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
;   TITLE OF INVENTION: PROTEIN
;   FILE REFERENCE: 07064-009002
;   CURRENT APPLICATION NUMBER: US/10/631,558
;   CURRENT FILING DATE: 2003-07-31
;   PRIOR APPLICATION NUMBER: US/09/940,235
;   PRIOR FILING DATE: 2002-04-09
;   PRIOR APPLICATION NUMBER: 09/471,349
;   PRIOR FILING DATE: 1999-12-23
;   PRIOR APPLICATION NUMBER: IN 3825/DEL/98
;   PRIOR FILING DATE: 1998-12-24
;   NUMBER OF SEQ ID NOS: 28
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 11
;   LENGTH: 1782
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Hybrid cassette
US-10-631-558-11

Query Match   98.2%;   Score 1222.6;   DB 53;   Length 1782;
Best Local Similarity 98.9%;   Pred. No. 0;
Matches 1231;   Conservative   0;   Mismatches   14;   Indels   0;   Gaps   0;

```
Qy     901 GATGTCAACACCAACGAATTGCTAAAAAGCGGAGCAGCTCTTAACAGCTAGGAACGTAAC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1438 GATGTCGATACCAACGAATTGCTAAAAAGTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGGCAACGTAAC 1497

Qy     961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAGCTAAAACTACTCACACAATCTC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1498 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAAGCTAAAACTACTCACACAATCTC 1557

Qy    1021 GATGCTTTGGTATTATGGACTATACCTTAACTGAAAAAGTAGAGGATAATCACGATGAC 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1558 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1617

Qy    1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1618 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGGAGAATGCTAGCTAT 1677

Qy    1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAAGTTTACAGCTACCTG 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGAACGAGAAGTTTACAGCTACCTG 1737

Qy    1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
           |||||||||||||||||||||||||||||||||||||||||||||
Db    1738 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1782
```

Db   1738 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA 1782

RESULT 12
US-09-569-920-2
; Sequence 2, Application US/09569920
; GENERAL INFORMATION:
; APPLICANT: Reddy, Vemuri B.
; APPLICANT: Lerner, Ethan
; TITLE OF INVENTION: VASODILATOR-THROMBOLYTIC FUSION PROTEIN
; TITLE OF INVENTION: AND CONJUGATES
; FILE REFERENCE: 10284-026001
; CURRENT APPLICATION NUMBER: US/09/569,920
; CURRENT FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 2
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
US-09-569-920-2

Query Match   98.1%;   Score 1221.2;   DB 24;   Length 1254;
Best Local Similarity 99.0%;   Pred. No. 0;
Matches 1229;   Conservative 0;   Mismatches 13;   Indels 0;   Gaps 0;

Qy   1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCGTCAACAACAGCCAATTAGTTGTT 66

Qy   61 AGCGTTGCTCGGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 AGCGTTGCTCGGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC 126

Qy   121 GACCTAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 GATCTAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 186

Qy   181 AAACCATTTGCTACTGATAGTGGCGCGGATGCCCACATAAACTTGAAAAAGCTGACTTACTA 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 AAACCATTTGCTACTGATAGTGGCGCGGATGTCACATAAACTTGAAAAAGCTGACTTACTA 246

Qy   241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 306

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 366

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCAACCGTCCAAGAATTTTTGCTAAGCGGACAT 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 AAAGATGGTTCGGTAACCTTGCCGACCAACCGTCCAAGAATTTTTGCTAAGCGGACAT 426

Qy   421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 486

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGTCTCAAA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGTCTCAAA 546

Qy   541 GATACTAAGCTATTGAAAACACTAGCTATCACATCTCAAGAATTACTA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 GATACTAAGCTATTGAAAACACTAGCTATCACATCTCAAGAATTACTA 606

Qy   601 GCTCAAGCACAAAGCATTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607 GCTCAAGCACAAAGCATTTAAACAAAACACCACCCAGGCTATACGATTTATGAACGTGAC 666

Qy   661 TCCTCAATCGTCACTCATGACAATGACATTTTTCCGTACGATTTTACCAATGACAAGAG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667 TCCTCAATCGTCACTCATGACAATGACATTTTTCCGTACGATTTTACCAATGACAAGAG 726

Db   658 GATCTAACATCACGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAATCA 717
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   181 AAACCATTTGCTACTGATAGTGGCGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   718 AAACCATTTGCTACTGATAGTGGCGCGGATGTCACATAAACTTGAGAAAGCTGACTTACTA 777

Qy   241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCAGTAACGACGACTACTTTGAGGTC 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   778 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCAGTAACGACGACTACTTTGAGGTC 837

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTCTACTTTGCTGAC 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   838 ATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTCTACTTTGCTGAC 897

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCAACCGTCCAAGAATTTTTGCTAAGCGGACAT 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   898 AAAGATGGTTCGGTAACCTTGCCGACCAACCGTCCAAGAATTTTTGCTAAGCGGACAT 957

Qy   421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   958 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 1017

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGTCTCAAA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1018 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGTCTCAAA 1077

Qy   541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACATCACATCTCAAGAATTACTA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1078 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACATCACATCTCAAGAATTACTA 1137

Qy   601 GCTCAAGCACAAAGCATTTAAACAAAACCCACCCAGGCTATACGATTATGAACGTGAC 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1138 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTATGAACGTGAC 1197

Qy   661 TCCTCAATCGTCACTCATGACAATGACATTTTTCCGTACGATTTTACCAATGACAAGAG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1198 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGACAAGAG 1257

Qy   721 TTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1258 TTTACTTACCGTGTTAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 1317

Qy   781 AATGAAGAAATAAACAACACTGAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1318 AATGAAGAAATAAACAACACTGAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1377

Qy   841 GAAAAGCCGTATGATCCCCTTTGATCGCAGTGCACTTGAACGTGTTCACCATCAAATACGTT 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1378 GAAAAGCCGTATGATCCCCTTTGATCGCAGTGCACTTGAACGTGTTCACCATCAAATACGTT 1437

Qy   901 GATGTCAACACCAACGAATTGCTAAAAGCGAGCAGCTCTTAACAGCTAGCGGAACGTAAC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1438 GATGTCGATACCAACGAATTGCTAAAAGTGCTAAAAGTAGAGCAGCTCTTAACAGCTAGCGGAACGTAAC 1497

Qy   961 TTGACTTCAGAGATTTATAGCATCGTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1498 TTGAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1557

Qy 1021 GATGCTTTTGGTATATGGACTATACCTTAAACTGAAAAAGTAGAGGATAATCACGATGAC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1558 GATGTCTTTTGGTATATGGACTATACCTTAAACTGAAAAAGTAGAGGATAATCACGATGAC 1617

Qy 1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAAGAAGTTTACAGCTATTATGAACGTAGCTAT 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1618 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAAGAAGTTTACAGCTATTATGAACGTAGCTAT 1677

Qy 1141 CATTTAGCCTATGATAAAGATGTTATACCGGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1678 CATTTAGCCTATGATAAAGATGTTATACCGGAAGAAGAACGAGAAGTTTACAGCTACCTG 1737

Qy 1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA 1245
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 1738 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA

RESULT 13
US-09-569-920-3
; Sequence 3, Application US/09569920
; GENERAL INFORMATION:
; APPLICANT: Reddy, Vemuri B.
; APPLICANT: Lerner, Ethan
; TITLE OF INVENTION: VASODILATOR-THROMBOLYTIC FUSION PROTEIN
; TITLE OF INVENTION: AND CONJUGATES
; FILE REFERENCE: 10284-026001
; CURRENT APPLICATION NUMBER: US/09/569,920
; CURRENT FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 8893
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-569-920-3

Query Match 98.1%; Score 1221.2; DB 24; Length 8893;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1229; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

RESULT 14
US-09-471-349-6
; Sequence 6, Application US/09471349
; GENERAL INFORMATION:
; APPLICANT: Sahni, Girish
; APPLICANT: Kumar, Rajesh
; APPLICANT: Roy, Chaiti
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE PROTEINS POSSESSING ALTERED
; TITLE OF INVENTION: PLASMINOGEN ACTIVATION CHARACTERISTICS AND A PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF SAID PROTEINS
; FILE REFERENCE: 07064/009001
; CURRENT APPLICATION NUMBER: US/09/471,349
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated primer
US-09-471-349-6

Query Match   97.2%;   Score 1209.8;   DB 20;   Length 1327;
Best Local Similarity   98.2%;   Pred. No. 0;
Matches 1223;   Conservative   0;   Mismatches   22;   Indels   0;   Gaps   0;

Qy   1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTTGTT 60
            ||||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   83 ATAGCTGGTTCCTGAATGGCTACTAGATCGTCCTTCGTAAATAACAGCCAATTGGTTGTT 142

Qy   61 AGGCTTGCTGGTACTGTTGAGGGGACGAGCAATCAAGACATTCAAGACAGCTTCTACTTGTCTACTTGTCTGAC 120
            ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   143 AGCGTTGCTGGTACTGTTGAGGGGACGAGCAATCAAGACAGCAATCAAGACAGCAATC 202

Qy   121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCAAGGTTAGTCCAAATCA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCAAGGCTTAAGTCCAAAATCA 262

Qy   181 AAACCATTTGCTACTGATAGTGGCGCGATGCGCCACATAAACTTGAAAAAGCTGACTACTA 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263 AAACCATTTGCTACTGATAGTGGCGCGATGTGGCGCGATGTCACATAAACTTGAGAAAAGCTGACTACTA 322

Qy   241 AAGGCTATTCAAGAACAATTGATCGCTACGTCCACAGTCAACGACGACTACTTTGAGGTC 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323 AAGGCTATTCAAGAACAATTGATCGCTACGTCCACAGTCAACGACGACTACTTTGAGGTC 382

Qy   301 ATTGATTTTGCAAGCGATGCAACCATTACTACTTGCTCTACTTTGCTGAC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   383 ATTGATTTTGCAAGCGATGCAACCATTACTACTTGCTCTACTTTGCTGAC 442

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   443 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTGCTAAGCGGACAT 502

Qy   421 GTGCGCGTTAGACCATATAAGAAAGAAAACAATACAAATCAAGCGAAATCGTTGATGTG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   503 GTGCGCGTTAGACCATATAAGAAAGAAAACAATACAAATCAAGCGAAATCGTTGATGTG 562

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGAGACCAGGTCTCAAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   563 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGAGACCAGGTCTCAAA 622

Qy   541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   623 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 682
Qy   601 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   683 GCTCAAGCACAAAGCATTTAAACAAAACCACCCAGGCTATACGATTTATGAACGTGAC 742

Qy   661 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTACCAATGGATCAAGAG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   743 TCCTCAATCGTCACTCATGACAATGACATTTCCGTACGATTTACCAATGGATCAAGAG 802

Qy   721 TTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGTCTG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   803 TTTACTTACCGTGTTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 862

Qy   781 AATGAAGAAATAAACAACACTGACCTGATCCTCGAGAAATATTACGTCCTAAAAAAGGG 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   863 AATGAAGAAATAAACAACACTGACCTGATCCTCTGAGAAATATTACGTCCTTAAAAAAGGG 922

Qy   841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTCGAAACTGTTCACCATCAAAATACGTT 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   923 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAAATACGTT 982

Qy   901 GATGTCAACACAACGAATTGCTAAAAAGGCAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   983 GATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1042

Qy   961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1043 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1102

Qy   1021 GATGCTTTTGGTATTCATAACCGTTATATGGGACTATACCTTAACTGGCAAGCGACCCGAGAAGGAGAGAATGCTAGCTAT 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1103 GATGCTTTTGGTATTCATAACCGTTATATGGGACTATACCTTAACTGGCAAGCGACCCGAGAAGGAGAGAATGCTAGCTAT 1162

Qy   1081 ACCAACCGTATCATAACCGTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1163 ACCAACCGTATCATAACCGTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1222

Qy   1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGATCGTTATACCGAAGAAGAAGTTTACAGCTACCTG 1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1223 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGATCGTTATACCGAAGAAGAAGTTTACAGCTACCTG 1282

Qy   1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1245
            ||||||||||||||||||||||||||||||||||||||||||||
Db   1283 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAA 1327

RESULT 15
US-09-940-235-6
; Sequence 6, Application US/09940235
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327

```
;  TYPE: DNA
;  ORGANISM: Streptococcus equisimilis
US-09-940-235-6

Query Match             97.2%;   Score 1209.8;   DB 39;   Length 1327;
Best Local Similarity   98.2%;   Pred. No. 0;
Matches 1223;   Conservative    0;   Mismatches   22;   Indels   0;   Gaps    0;

Qy      1 ATTGCTGGACCTGAGTGGTCTGCTAGACCGTCCATCTGTCAACACAGCCAATTAGTTGTT 60
          ||||||||| ||||||||||||||| ||||||| ||||| |||||||||||||||||||||
Db     83 ATAGCTGGTCCTGAATCGTCCTACTAGATCGTCCTTCTGTAAATAAACAGCCAATTGGTTGTT 142

Qy     61 AGCGGTTGCTGTCGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 AGCGGTTGCTGTCGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTGAAATC 202

Qy    121 GACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGAGCAAGCAAGCTTAAGTCCAAAATCA 180
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    203 GATCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGAGCAAGCAAGCTTAAGTCCAAAATCA 262

Qy    181 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 AAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAGAAAAGCTGACTTACTA 322

Qy    241 AAGGCTATTCAAGAACAATTGATCGTCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 AAGGCTATTCAAGAACAATTGATCGTCTAACGTCCACAGTAACGACGACTACTTTGAGGTC 382

Qy    301 ATTGATTTTGCAAGCGATGCAACCATTACTAGTCGAAACGGCAAGGTCTCTACTTTGCTGAC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 ATTGATTTTGCAAGCGATGCAACCATTACTAGTCGAAACGGCAAGGTCTCTACTTTGCTGAC 442

Qy    361 AAAGATGGTTCGGTAACCTGCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 AAAGATGGTTCGGTAACCTGCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGACAT 502

Qy    421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 562

Qy    481 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTCAGACCAGGTCTTCAAA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    563 GAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTCAGACCAGGTCTTCAAA 622

Qy    541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA 682

Qy    601 GCTCAAGCACAAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTATGAAACGTGAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 GCTCAAGCACACAAAGCATTTTAAACAAAACACCCAGGCTATACGATTTATGAAACGTGAC 742

Qy    661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAG 802

Qy    721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    803 TTTACTTACCGTGTCTAAAAATCGGGAACAAGCTTATAGGATCAATAAAAATCTGGTCTG 862

Qy    781 AATGAAGAAATAAACAACACTGACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    863 AATGAAGAAATAAACACACTGACCTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 922

Qy    841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    923 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATACGTT 982

Qy    901 GATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGTCTTTAACAGCTAGCGAACGTAAC 960
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    983 GATGTCGATACCAACCAACGAATTGCTAAAAAGTGAGCAGTCTTTAACAGCTAGCGAACGTAAC 1042

Qy    961 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1043 TTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAATCTC 1102

Qy   1021 GATGCTTTTGGTATTATGGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1103 GATGCTTTTGGTATTATGGGACTATACCTTAACTGGAAAAGTAGAGGATAATCACGATGAC 1162

Qy   1081 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAAGAATGCTAGCTAT 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1163 ACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAAGAATGCTAGCTAT 1222

Qy   1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAACGAGAAGTTTACAGCTACCTG 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1223 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAACGAGAAGTTTACAGCTACCTG 1282

Qy   1201 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAATAA 1245
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1283 CGTTATACAGGGACACCTATACCTGATAACCCTAACGACACAAATAA 1327


Search completed: May  6, 2004, 07:44:45
Job time : 4633 secs
```

```
          GenCore version 5.1.6
     Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      May 6, 2004, 03:28:18 ; Search time 290 Seconds
                                (without alignments)
                         3661.179 Million cell updates/sec

Title:           US-09-882-509-3
Perfect score:   1245
Sequence:        1 attgctggacctgagtggct.........ataaccctaacgacaaata 1245
Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched:        1493552 seqs, 426402401 residues

Total number of hits satisfying chosen parameters:    2987104

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending Patents NA New:*
                 1: /cgn2_6/ptodata71/pna/PCT_NEW_COMB.seq:*
                 2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
                 3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
                 4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
                 5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
                 6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
                 7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	41.6	3.3	1293	7	US-60-551-161-131	Sequence 131, App
2	39.2	3.1	99971	6	US-10-767-471-10561	Sequence 10561, App
3	38.2	3.1	2232	6	US-10-808-727-45	Sequence 45, Appl
4	37.4	3.0	110608	6	US-10-775-169-193	Sequence 193, App
5	36.4	2.9	753	5	US-09-570-737A-647	Sequence 647, App
c 6	35.8	2.9	765267	7	US-10-767-471-10616	Sequence 10616, A
7	35.4	2.8	2160	7	US-60-551-121-143	Sequence 143, App
8	35.2	2.8	1184710	6	US-10-796-280-12394	Sequence 12394, A
c 9	35	2.8	1252	6	US-10-786-892-69	Sequence 69, Appl
10	35	2.8	1790242	6	US-10-767-471-10805	Sequence 10805, A
c 11	34.8	2.8	1984	6	US-10-100-683-128	Sequence 128, App
12	34	2.7	2427	5	US-09-537-710D-5	Sequence 5, Appli
c 13	33.8	2.7	46143	6	US-10-796-280-12192	Sequence 12192, A
14	33.8	2.7	185248	6	US-10-796-280-12228	Sequence 12228, A
c 15	33.6	2.7	65385	6	US-10-796-307-8817	Sequence 8817, Ap
16	33.4	2.7	837	6	US-10-488-619-1789	Sequence 1789, Ap
17	33.4	2.7	158417	6	US-10-796-307-8778	Sequence 8778, Ap
18	33.4	2.7	158417	6	US-10-803-180-1675	Sequence 1675, Ap
c 19	33.2	2.7	65459	7	US-60-563-440-12039	Sequence 12039, A
20	33	2.7	276	6	US-10-724-972A-250	Sequence 250, App
21	33	2.7	82553	6	US-10-100-683-12035	Sequence 12035, A
22	33	2.7	87394	7	PCT-US04-09319-6	Sequence 6, Appli
c 23	32.8	2.6	201	6	US-10-796-280-21377	Sequence 21377, A
c 24	32.8	2.6	201	6	US-10-796-280-21644	Sequence 21644, A
c 25	32.8	2.6	142613	6	US-10-796-307-8749	Sequence 8749, Ap
c 26	32.8	2.6	154750	7	US-60-563-440-12165	Sequence 12165, A
c 27	32.8	2.6	663098	6	US-10-796-280-12289	Sequence 12289, A
28	32.6	2.6	2610	7	US-10-784-480-243	Sequence 243, App
29	32.6	2.6	3573	7	US-10-559-949-798	Sequence 798, App
c 30	32.6	2.6	31127	6	US-10-100-683-12644	Sequence 12644, A
31	32.6	2.6	34655	6	US-10-796-280-12337	Sequence 12337, A
c 32	32.6	2.6	95990	7	US-60-563-440-11855	Sequence 11855, A
c 33	32.6	2.6	163708	6	US-10-796-280-12588	Sequence 12588, A
34	32.6	2.6	546025	6	US-10-767-471-10717	Sequence 10717, A
c 35	32.4	2.6	665	6	US-10-767-701-21191	Sequence 21191, A
c 36	32.4	2.6	643250	6	US-10-796-280-12230	Sequence 12230, A
37	32.2	2.6	201	6	US-60-550-051-6793	Sequence 6793, Ap
38	32.2	2.6	359	5	US-09-796-692B-4345	Sequence 4345, Ap
39	32.2	2.6	966	6	US-10-767-701-6462	Sequence 6462, Ap
40	32.2	2.6	61009	7	US-60-563-440-11881	Sequence 11881, A
41	32.2	2.6	816609	7	US-60-550-051-2987	Sequence 2987, Ap
42	32	2.6	201	6	US-10-796-280-49577	Sequence 49577, A
43	32	2.6	201	6	US-10-796-280-65927	Sequence 65927, A
44	32	2.6	477	6	US-10-767-701-28416	Sequence 28416, A
45	32	2.6	531	6	US-10-767-701-29142	Sequence 29142, A

ALIGNMENTS

```
RESULT 1
US-60-551-161-131
; Sequence 131, Application US/6051161
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Alterman, Eric
; APPLICANT: Cano, Raul J.
; APPLICANT: Hamrick, Alice
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Stress-Related Proteins and Uses
; TITLE OF INVENTION: Therefore
; FILE REFERENCE: 043556/261831
; CURRENT APPLICATION NUMBER: US/60/551,161
; CURRENT FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 131
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Lactobacillus acidophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1293)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: ORF 1295; FtsY - Cell division protein
US-60-551-161-131

Query Match             3.3%;  Score 41.6; DB 7; Length 1293;
Best Local Similarity   51.0%; Pred. No. 0.074;
Matches    98; Conservative   0; Mismatches   94; Indels   0; Gaps   0;

Qy      437 ATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTGGAATATACTGTACAGT 496
                |||||||  ||  | ||   || || | |  ||||  ||| ||  ||| | || ||
Db       59 AAAAAGATGAAGAAAACAAAGTCAAGAACAGGAAGATATAGAAGAAACAACTGATATGG 118

Qy      497 TTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAAGATACTAAGCTATTGA 556
                ||   |||   || | ||| || || ||     ||| ||||   | ||  |||
Db      119 AAAATGACTCTTCTAATGTTGAAGATTCAGAAAAGTAGATAATGATACTTCAAGATTCGG 178

Qy      557 AAACACTAGTCATCGGTGACACCATCACATCTCAAGAATTACTAGCTCAAGCACAAAGCA 616
                ||||||  ||| || ||  | |   |  ||||| ||| |||| |  ||| |  |
Db      179 AAAAAATAGAAGAAACTGAAGAAAAACACTTCTGATGAATCATCGGTTGAAGAAGAATTA 238

Qy      617 TTTTAAACAAAA 628
                | |||   ||
Db      239 GTGAAAATGATA 250
```

```
Qy   723  TACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAAATCTGGTCTGAA  782
          |||  |||  |  ||||||      |||  || |  || |  | || |   |||||||
Db  1620  TAATCATCCTGAAAATTATTTTAAAGAAGAAGATATTATTATGATGAAAATGATGATATGGA 1679

Qy   783  TGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATT---                  842
          |  || |||| |       ||| |||  | |  | |||
Db  1680  AGTAAAAGTTAAAAAAATAGGTGTCACATTAAAAAAAATTTGAACCACTTAAAAAATGG--- 1736

Qy   843  AAAGCCGTATGATCGATCCCTTGATGCGCAGTCACTGAAACTGTTCACCATCAAATACGTTGA 902
          |||     || ||    |  || || |  |  |  | |   |  ||  |   ||| |
Db  1737  AAATGTTAGTGAAACCATTAAATTGATTCATTTAGGAAATAAGATAAAAARCACACATTGA  1796

Qy   903  TGTCAACACCAACGAATTGCTAAAAAGCGGAGCAGCTTCTTAACAGCTA             949
          | |  | | ||| ||  |  |    ||  ||  || || |  |  |||||
Db  1797  AGCTATAAACAACGATATTCAAATTATTAAACAAGAATTACAAGCTA              1843
```

RESULT 2
US-10-767-471-10561/c
; Sequence 10561, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10561
; LENGTH: 99971
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(99971)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-767-471-10561

Query Match   3.1%; Score 39.2; DB 6; Length 99971;
Best Local Similarity   48.2%; Pred. No. 2.1;
Matches 107; Conservative 1; Mismatches 114; Indels 0; Gaps 0;

```
Qy   674  CTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCATG    733
Db 42609  CTTAAATCAGTGAAACTTACCACACATTATTAGATTAAAGAGAAAATATCATATGACCATTT 42550

Qy   734  TCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAAATCTGGTCTGAATGAAGAAATAA   793
Db 42549  AAAAAGATTCAGGACAAGTATTTRATAAAATTCAACATTTATGTTTGATTAAAAAGTAAA   42490

Qy   794  ACAACACTGACCTGACCTCTCGAGAAATATTACGTCCTTAAAAAAGGGGAAAAGCCGTATG   853
Db 42489  TAACAAAACCTCTTAGCAAGCTTAGAGGATACTTCCATAAATCGGAAAAGGATACTTCTA   42430

Qy   854  ATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAAT                     895
Db 42429  AACATCCTATGCAAATAACATGAAACCAGTCACTCACTGAGAT                   42388
```

RESULT 3
US-10-808-727-45
; Sequence 45, Application US/10808727
; GENERAL INFORMATION:
; APPLICANT: Oh, Steven
; APPLICANT: Chishti, Athar
; APPLICANT: Liu, David
; APPLICANT: Li, Xuerong
; APPLICANT: Goel, Vikas
; TITLE OF INVENTION: Band 3-Binding Malaria Polypeptides and Uses Thereof
; FILE REFERENCE: S1237.7019US01
; CURRENT APPLICATION NUMBER: US/10/808,727
; CURRENT FILING DATE: 2004-03-25
; PRIOR APPLICATION NUMBER: US 10/087,464
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/272,930
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
; LENGTH: 2232
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
US-10-808-727-45

Query Match   3.1%; Score 38.2; DB 6; Length 2232;
Best Local Similarity   51.1%; Pred. No. 0.86;
Matches 116; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

RESULT 4
US-10-775-169-193/c
; Sequence 193, Application US/10775169
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dorner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 193
; LENGTH: 110608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-193

Query Match   3.0%; Score 37.4; DB 6; Length 110608;
Best Local Similarity   52.2%; Pred. No. 7.3;
Matches   83; Conservative 0; Mismatches   76; Indels 0; Gaps 0;

```
Qy   415  GGACATGTGCGCGTTAGCACATATAAGAGAAAAACCAATACAAATACAAGCGAAATCTGTT   474
Db 46041  GGGGATGGGAGCGGAGCGAGTCCTTCTCAAAAAAAATAAAATAAAACAAAAAACAAATGAGAA 45982

Qy   475  GATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGT   534
Db 45981  ACTGTAAAATAAACCGTAAACTGTGTGAAATAGGTATTTAGGGAAATCTCCACTAGAAGT   45922

Qy   535  CTCAAAGATACTAAGCTATTGAAAAACACTAGCTATCCGT                       573
Db 45921  CTCTGAGTTTGAAGGTTTTGAAATTGAGTGCTTTCTGT                        45883
```

RESULT 5
US-09-570-737A-647/c
; Sequence 647, Application US/09570737A
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Helper, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C14
; CURRENT APPLICATION NUMBER: US/09/570,737A
; CURRENT FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 09/568,100
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: 09/536,857
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 09/483,672
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 09/443,686
; PRIOR FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 09/439,313
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 09/352,616
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/288,946
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/232,149
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 09-159,812
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 09/115,453
; PRIOR FILING DATE: 1998-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 780
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 647
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(753)
; OTHER INFORMATION: n = A,T,C or G
US-09-570-737A-647

Query Match   2.9%; Score 36.4; DB 5; Length 753;
Best Local Similarity 50.0%; Pred. No. 1.8;
Matches   88; Conservative   0; Mismatches 88; Indels   0; Gaps   0;

Qy   672 CACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCA 731
          |||||||||| ||||| ||||||||||||||||| ||||||||
Db   658 CATTAATGTCAGAGATATTCCCAGTATAATTTTCNTAAGACATGTAATCTTCACTCACCA 599

Qy   732 TGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTGAATCGAAGAAAT 791
          ||||| ||||||||
Db   598 CTTATCACAGTAGGATGACTCTTAAGTCTNGTATTAACAACCTGACTACATTAAACAAAG 539

Qy   792 AACAACAACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGGGAAAAGC 847
          |||||||||||||||||||||||||
Db   538 CAGAGCCCATGAATTGATACAAGGTGGATAAGCCTTCATCACATGAACAAAACACC 483

RESULT 6
US-10-767-471-10616/c
; Sequence 10616, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10616
; LENGTH: 765267
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(765267)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10616

Query Match   2.9%; Score 35.8; DB 6; Length 765267;
Best Local Similarity 55.1%; Pred. No. 45;
Matches   70; Conservative   0; Mismatches 57; Indels   0; Gaps   0;

Qy   744 GGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTGAATGAGAAATAAACAACACTGA 803
          ||| ||||| |||||||||||||||| ||||||
Db 210753 GAAATGATCCTAAAATATGTAAAATACATGTTTTTCTAAATAAAGATATAAGCAACCAGTA 210694

Qy   804 CCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGGGAAAAGCCTATGATCCCTTTGA 863
          ||||||||||||||||||| ||||||
Db 210693 GAAGAGCTATAATATTTGTTTGAACATTAAAAAAGGAAAGATTATAATCATAGTTATCA 210634

Qy   864 TCGCAGT 870
          || ||||
Db 210633 TCACAGT 210627

RESULT 7
US-60-551-121-143
; Sequence 143, Application US/60551121
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Alterman, Eric
; APPLICANT: Cano, Raul J.
; APPLICANT: Hamrick, Alice
; TITLE OF INVENTION: Lactobacillus acidophilus Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Carbohydrate Utilization-Related Proteins and Uses
; TITLE OF INVENTION: Therefore
; FILE REFERENCE: 043556/264779
; CURRENT APPLICATION NUMBER: US/60/551,121
; CURRENT FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 143
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Lactobacillus acidophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2160)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: ORF 1796; plnG
US-60-551-121-143

Query Match   2.8%; Score 35.4; DB 7; Length 2160;
Best Local Similarity 48.3%; Pred. No. 5.4;
Matches   99; Conservative   0; Mismatches 106; Indels   0; Gaps   0;

Qy   85 ACGAATCAAGACATTAGTCTTAAATTTTTGAAATTGACCTAACATCACGACCTGCTCAT 144
          |||| ||| ||| ||| ||||||||||||| ||||
Db   82 AAGAAATATCATTCTAGAGTATCATTAGCACGTATCATTAGCACGTACTAATTTA 141

Qy   145 GGAGGAAAGACACAGACCAAGGCTAAGTCCAAAATCAAAACCATTTGCTACTGATAGTGGC 204
          |||||||||| |||| ||||||||||
Db   142 GAAGGAACTACCGCCCTTGGCTTAGTCAAAACAGCACAAACATTTAATTTGAAAACTGAA 201

Qy   205 GCGATGCCACATAAACTTGAAAAAGCTGACTTACTAGAAGCTATTCAAGAACAATTGATC 264
          |||| |||||| ||| ||||||||
Db   202 GCGGTCAAGCTGATATGTCCTTGTTTGATCCAGATACGGATATCCAATATCCTTTTATT 261

Qy   265 GCTAACGTCCACAGTAACGACGACT 289
          ||| || |||||||||
Db   262 GTTCATGTCTAAAACAAGGCGAAT 286

```
Qy    746 AACAAGCTTATGAGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACC 805
              ||||||||||||| |||  |||| ||||  ||||| |||||||||||||||||||
Db    440 AATACCCTAATAACCCAGATAAAACCCCAATTTTAAAAAGTTTTTTAAACTTTTACGCC 381

Qy    806 TGATCTCTGAGAAATATTAGTCCTTAAAAAAGGGAAAAGCCGTATGATCCCTTTGATC 865
              |||||||||||||||||||  ||||||||||||||||||||||||| |||||||
Db    380 TTGAATGCAAAAATCTATTTTTCTATGGAATAGGAAAACCCGTAAAAACTTAAAGAAA 321

Qy    866 GCAGTCACTGAAACTGTTCACCACAAATACGTTGATGTCAACCAACGAATTGCTA 924
              ||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    320 GGATAAAATAATAAAGACACAAACGAACACGCTGAAGTCAAAACAAACAAATGGCTA 262
```

RESULT 10

US-10-767-471-10805

; Sequence 10805, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10805
; LENGTH: 1790242
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1790242)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10805

```
Query Match             2.8%;  Score 35;  DB 6;  Length 1790242;
Best Local Similarity  48.6%;  Pred. No. 99;
Matches  122;  Conservative   1;  Mismatches  126;  Indels   2;  Gaps   1;

Qy    249 TCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAGGTCATTGATT 308
              |||||||||: || |  ||||  ||| || |||| |||||||||||||||||||||
Db 545236 TAAACARCAAGTAAAGAAATAATCTAAAGCAATTGAAAATATTTGACTAATGAAAT 545295

Qy    309 TGCAAGCGATGCAACCATTACTGATCGAAAACGGCAAGGTCTACTTTGCTGACAAAGATGG 368
              ||||| |||||| ||||| ||||  |||||||||| |||||||  ||| |||| ||
Db 545296 TGAAAACAAAGCACACCAAGCTCATGGAA--TGCAGCAAAACCAGTACTAAGAGGGAAGT 545353

Qy    369 TTCGGTAACCTTGCCGACCCAACCTGTCCAAGAAATTTTTGCTAAGCGGACATGTGCGCGT 428
              |||  ||||| || || ||||||  |||  |||  ||||  ||||| |  ||||
Db 545354 TTCCATGACTGTTGACATTCAAAAAGAAGAAGAGATTTAAGTAAATAATAATCAACTTTGCA 545413

Qy    429 TAGACCATATAAAGAAGAAAAACCAATACAAAATCAAGCGAAAATCTGTTGATGTGGAATATAC 488
              ||  |||||||| ||||||  |  |||||||||||||  |||||| ||||||||||||
Db 545414 TCTCAAATAAGTAGAAAAAAGAAGAACAAATTAAGCCCAAAGTTAGTAGAAGGAAAGAAAT 545473

Qy    489 TGTACAGTTTA 499
              |||||| |||
Db 545474 GGTAAAGATTA 545484
```

RESULT 11

US-10-100-683-128/c

; Sequence 128, Application US/10100683
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS900
; CURRENT APPLICATION NUMBER: US/10/100,683
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 60/043,576

RESULT 8

US-10-796-280-12394

; Sequence 12394, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001510
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12394
; LENGTH: 1184710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1184710)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-796-280-12394

```
Query Match             2.8%;  Score 35.2;  DB 6;  Length 1184710;
Best Local Similarity  53.7%;  Pred. No. 77;
Matches   73;  Conservative   0;  Mismatches  63;  Indels   0;  Gaps   0;

Qy     51 ATTAGTTGTTGAGCGTTGCTGGTTACTGTTGAGGGACGAATCAAGACATTAGTAGTCTTAAATT 110
              |||||||||| |  | |||| |   ||||||||  |||  |||  |||||||| |||
Db 790383 ATTGTTGGTTTCCACAGATGCTGATGCTGAGGCAAGGCAAGAAATCAAGTCAAAAT TGGTTTATT 790442

Qy    111 TTTTGAAATTGACCTAACATCACGACACCTGCTCATCGGAGAAAAGACAGAGCAAGGCTTAAG 170
              ||  |  ||||  ||||   |||||||||||| |||  |||| || |||||| ||||
Db 790443 TTGGGAGGTACAGGAAACACCAGTAGAAAGTGAGGAAATAAGACAGGGASGGGASGGAAGGCA 790502

Qy    171 TCCAAAATCAAAACCA 186
              |||||||||||||||
Db 790503 GCCAATAACGTAACCA 790518
```

RESULT 9

US-10-786-892-69/c

; Sequence 69, Application US/10786892
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/041003
; CURRENT APPLICATION NUMBER: US/10/786,892
; CURRENT FILING DATE: 2004-02-25
; PRIOR APPLICATION NUMBER: US 09/881,752
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/833,457
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 1252
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(1198)
US-10-786-892-69

```
Query Match             2.8%;  Score 35;  DB 6;  Length 1252;
Best Local Similarity  49.7%;  Pred. No. 5.6;
Matches   89;  Conservative   0;  Mismatches  90;  Indels   0;  Gaps   0;
```

```
;     OTHER INFORMATION: n= a or g or c or t/u
US-09-537-710D-5

Query Match                2.7%;   Score 34;   DB 5;   Length 2427;
Best Local Similarity     50.6%;   Pred. No. 14;
Matches    82;  Conservative    0;  Mismatches    80;  Indels    0;  Gaps    0;

Qy      205  GGGATGCCACATAAACTTGAAAAAGTCGACTTACTAAAAGGCTATTCAAGAACAATTGATC  264
             ||||| ||||  |  |||||||||  || ||||||  |||| |||||||||||||||||||
Db     1357  GCTGCGCCATCTGAGATTAATAATATTGATTTTCGAGGTGTCTCAAAGGTCAGAGTATC  1416

Qy      265  GCTAACGTCCACAGTAACGACGACTACTTTGAGGTCATTGATTTTGCAAGCGGATGCAACC  324
             |  |||||||| || |||| || || |||| ||||||| || |||  ||  |||||||| |
Db     1417  CCAAATCACACCTGTCGTGACGTGTGGACAGAGTACCATGACATGGGAATTGCTGGGATC  1476

Qy      325  ATTACTGATCGAAACGGCAAGGTCACTTTGCTGACAAAGAT  366
             ||  |||||||||||  ||| |  ||||||||||||||  |
Db     1477  AAAGCTATCCGCTGAGTATAAGGTCTACACTGCTGGTGAAGCT  1518

RESULT 13
US-10-796-280-12192/c
;   Sequence 12192, Application US/10796280
;   GENERAL INFORMATION:
;     APPLICANT: CARGILL, Michele et al.
;     TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;     TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
;     FILE REFERENCE: CL001510
;     CURRENT APPLICATION NUMBER: US/10/796,280
;     CURRENT FILING DATE: 2004-03-10
;     NUMBER OF SEQ ID NOS: 68533
;     SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 12192
;     LENGTH: 46143
;     TYPE: DNA
;     ORGANISM: Homo sapiens
US-10-796-280-12192

Query Match                2.7%;   Score 33.8;   DB 6;   Length 46143;
Best Local Similarity     48.2%;   Pred. No. 54;
Matches    95;  Conservative    0;  Mismatches   102;  Indels    0;  Gaps    0;

Qy      453  ACAAAATCAAGCGAAATCGTTGATCGTGGAATATACTGTACAGTTTACTCCCTTAAACCC  512
             ||||  |||| ||||| ||| ||| || ||| ||||||| || |||||| ||| ||||||
Db    19113  AGAGAATTTTGAGAGTTATTCTGACATTCAATTGTAAACAGGGCAATTTGTAAACAAT  19054

Qy      513  TGATGACGATTTCAGACCAGTCTCAAAGATACTAAGCTATTGAAAAACACTAGCTATCGG  572
             ||  || || ||||| || || ||||| ||| |||| |||||||  ||| |||| ||||
Db    19053  TGCAAAGAATTGGTAAACAATTGCAAGCATTTTAAACAAACCCA  18994

Qy      573  TGACACCATCACATCTCAAGAATTACTAGCTCAAGCACAAGCATTTTAAACAAACCCA  632
             ||||||||| ||||  |||| ||||| ||| |||| |  ||||| ||||||| |||||
Db    18993  ACTTGTAAACAATTGCAAACAATTTGTAAATAAGTGTAAATAAGTTGTAAACAATTACAAACAAAGCCCA  18934

Qy      633  CCCAGGCTATACGATTT  649
             ||  |||  ||||| ||
Db    18933  GAGAGGGGAGTGAATT  18917

RESULT 14
US-10-796-280-12228
;   Sequence 12228, Application US/10796280
;   GENERAL INFORMATION:
;     APPLICANT: CARGILL, Michele et al.
;     TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;     TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
;     FILE REFERENCE: CL001510
;     CURRENT APPLICATION NUMBER: US/10/796,280
;     CURRENT FILING DATE: 2004-03-10
;     NUMBER OF SEQ ID NOS: 68533
;     SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 12228
;     LENGTH: 185248
```

```
;     PRIOR FILING DATE: 1997-04-11
;     PRIOR APPLICATION NUMBER: US 60/047,601
;     PRIOR FILING DATE: 1997-05-23
;     PRIOR APPLICATION NUMBER: US 60/056,845
;     PRIOR FILING DATE: 1997-08-22
;     PRIOR APPLICATION NUMBER: US 60/043,580
;     PRIOR FILING DATE: 1997-04-11
;     PRIOR APPLICATION NUMBER: US 60/047,599
;     PRIOR FILING DATE: 1997-05-23
;     PRIOR APPLICATION NUMBER: US 60/056,664
;     PRIOR FILING DATE: 1997-08-22
;     PRIOR APPLICATION NUMBER: US 60/043,314
;     PRIOR FILING DATE: 1997-04-11
;     PRIOR APPLICATION NUMBER: US 60/047,632
;     PRIOR FILING DATE: 1997-05-23
;     PRIOR APPLICATION NUMBER: US 60/056,892
;     PRIOR FILING DATE: 1997-08-22
;   Remaining Prior Application data removed - See File Wrapper or PALM.
;     NUMBER OF SEQ ID NOS: 13368
;     SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 128
;     LENGTH: 1984
;     TYPE: DNA
;     ORGANISM: Homo sapiens
US-10-100-683-128

Query Match                2.8%;   Score 34.8;   DB 6;   Length 1984;
Best Local Similarity     48.1%;   Pred. No. 7.7;
Matches    99;  Conservative    0;  Mismatches   107;  Indels    0;  Gaps    0;

Qy      522  TTTCAGACCAGGTCTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCAT  581
             |||  ||  |||||||||  ||||| || ||||||  ||| || |||  ||  |||| ||
Db     1714  TTTAAAACTGAATGCAAAATTTACTGGGTTACAGTAGTATACAATGATAAACTCTGACAACAA  1655

Qy      582  CACATCTCAAGAATTACTAGCTCAAGCACAAAGACATTTTAAACAAAACCCACCCAGGCTA  641
             |||||||||| ||||  ||| |||| |  ||| || |||| |||| || ||  || |||
Db     1654  AAAAATTTAAATATAATATAATAATTTCTTTTCAAACATAAAGTTCACATACTGCAAACAGAAGA  1595

Qy      642  TACGATTTATGAGACGTGACTCCTCCAATCGTCACTCACTGACAAATGACACATTTTCCGTACGAT  701
             ||  ||||||  ||||||||| || || ||  || |||| |  |||||  || || ||||| |||
Db     1594  CACTGTGTGATTTCATTCCTAGTAAGGAGTTAGTCTTGCATATGATATAGGCACTCAATATT  1535

Qy      702  TTTACCAAATGGATCAAGAGTTTACTT  727
             |||||  || |  || ||||| ||||
Db     1534  TGTTCAACTGGACCTAAAATTTACAT  1509

RESULT 12
US-09-537-710D-5
;   Sequence 5, Application US/09537710D
;   GENERAL INFORMATION:
;     APPLICANT: Dahlquist, Anders
;     APPLICANT: Stahl, Ulf
;     APPLICANT: Lenman, Marit
;     APPLICANT: Banas, Antoni
;     APPLICANT: Ronne, Hans
;     TITLE OF INVENTION: A new class of enzymes in the biosynthetic pathway for the produc
;     TITLE OF INVENTION: of
;     TITLE OF INVENTION: triacylglycerol and recombinant DNA molecules encoding these enz
;     FILE REFERENCE: BASFnae337799PCT1-15
;     CURRENT APPLICATION NUMBER: US/09/537,710D
;     CURRENT FILING DATE: 2000-03-30
;     PRIOR APPLICATION NUMBER: EP 99106656.4
;     PRIOR FILING DATE: 1999-04-01
;     NUMBER OF SEQ ID NOS: 21
;     SOFTWARE: WordPerfect version 6.1
;   SEQ ID NO 5
;     LENGTH: 2427
;     TYPE: DNA
;     ORGANISM: Arabidopsis thaliana
;     FEATURE:
;     NAME/KEY: unsure
;     LOCATION: 1...2427
```

Thu May   6 12:30:22 2004

us-09-882-509-3.rnpn

Page 6

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-796-280-12228

Query Match   2.7%;   Score 33.8; DB 6; Length 185248;
Best Local Similarity 48.2%; Pred. No. 95;
Matches   95; Conservative 0; Mismatches 102; Indels   0; Gaps   0;

Qy   453 ACAAAATCAAGCGAAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTTAAACCC 512
              | |||| ||||| |||| | ||| | ||||| | |||| ||| | ||| |||
Db 47472 AGAGAATTTTGAGAGTTATTCTGACATTCAATTGTAAACAAGGGCAATTTGTAAACAAT 47531

Qy   513 TGATGCACGATTTCAGACCAGGTCTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGG 572
              | | || |||| || ||| | || ||||| || | | |||| |||| || | |
Db 47532 TGCAAAGAATTGGTAAACAATTGCAAACGATTGTAAACAATTGCAAACAATGTGTAAACA 47591

Qy   573 TGACACCATCACATCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCA 632
              | ||| | | ||| | ||| | ||| | | || |||| |||||| ||||
Db 47592 ACTTGTAAACAATTGCAAACAATTTGTAAATAAGTTGTAAACAATTACAAACAAAGCCCA 47651

Qy   633 CCCAGGCTATACGATTT 649
              |||| || ||| |||
Db 47652 GAGAGGGGAAGTGAATT 47668

RESULT 15
US-10-796-307-8817/c
; Sequence 8817, Application US/10796307
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001509
; CURRENT APPLICATION NUMBER: US/10/796,307
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 44201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8817
;   LENGTH: 65385
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-796-307-8817

Query Match   2.7%;   Score 33.6; DB 6; Length 65385;
Best Local Similarity 59.4%; Pred. No. 71;
Matches   57; Conservative 0; Mismatches 39; Indels   0; Gaps   0;

Qy   718 GAGTTTACTTACCCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGT 777
              ||| ||||| || || || || ||| || | | | | ||| ||||| |||
Db 12224 GGGAATACTGATGAAGTTTCCAAGAGGAAATGAGGTTTCACATTAATAACAAGTTTAGC 12165

Qy   778 CTGAATGAAGAAATAAACAACAACACTGACCTGATCTCT 813
              || | | |||| ||| | || ||| | || ||| ||||
Db 12164 CTGCAAGAAACACAAAATAGTACTCCTCTTATCTCT 12129

Search completed: May   6, 2004, 06:25:24
Job time : 295 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 6, 2004, 01:11:57 ; Search time 3628 Seconds
(without alignments)
10247.634 Million cell updates/sec

Title: US-09-882-509-3
Perfect score: 1245
Sequence: 1 attgctggacctgagtggct.........ataaccctaacgacaaataa 1245

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 1491109276 residues

Total number of hits satisfying chosen parameters: 5502578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: em_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	51.6	4.1	712	13	BX416727	BX416727 BX416727
c 2	49.4	4.0	1124	13	BX436282	BX436282 BX436282
c 3	49.4	4.0	1217	13	BQ678748	BQ678748 AGENCOURT
c 4	48	3.9	1175	13	BX423299	BX423299 BX423299

	No.						ID	Description
c	5	47.8	3.8	1101	29	CNS0100X	AL098379 Drosophil
c	6	47	3.8	994	13	BX414650	BX414650
c	7	46.4	3.7	1101	29	CNS0039G	AL063921 Drosophil
c	8	46.2	3.7	885	13	BX425603	BX425603
c	9	45.4	3.6	1085	13	BX414831	BX414831
c	10	45.4	3.6	1201	13	BX361080	BX361080
c	11	44.4	3.6	1007	29	CNS06X9S	AL419462 T3 end of
c	12	44.2	3.6	1101	29	CNS0LT2	AL078714 Drosophil
c	13	44.2	3.6	878	29	CNS0187R	AL108993 Drosophil
c	14	44.2	3.6	1101	29	CNS0HE9	AL073374 Drosophil
c	15	44.2	3.6	1109	13	BX375784	BX375784
c	16	42.8	3.4	1101	29	CNS0LQO	AL068607 Drosophil
c	17	42.4	3.4	1180	13	BX436369	BX436369
c	18	42	3.4	1061	13	BX414576	BX414576
c	19	41.8	3.4	500	12	BP188195	BP188195
c	20	41.8	3.4	1201	13	BX419736	BX419736
c	21	41.6	3.3	990	13	BX382345	BX382345
c	22	41.6	3.3	1081	28	CC238890	CC238890
c	23	41	3.3	796	12	BM398641	BM398641 CH261-80N
c	24	40.8	3.3	508	28	AQ591034	AQ591034 5009-0-48
c	25	40.8	3.3	678	28	BH473995	BH473995 HS_5391_B
c	26	40.8	3.3	741	14	CF134481	CF134481 BOGWU69TR
c	27	40.8	3.3	976	29	AG136173	AG136173 CFG12-85.
c	28	40.8	3.3	1101	29	CNS000D1	AL065414 Drosophil
c	29	40.8	3.3	1103	13	BX387723	BX387723
c	30	40.8	3.3	1201	13	BX394291	BX394291
c	31	40.6	3.3	507	12	BM879961	BM879961 ku03c03.Y
c	32	40.6	3.3	996	29	CNS00FUH	AL071063 Drosophil
c	33	40.6	3.3	1091	13	BX424950	BX424950
c	34	40.2	3.2	337	29	AL765028	AL765028 Arabidops
c	35	40.2	3.2	500	9	AU088371	AU088371 AU088371
c	36	40	3.2	649	9	AV733574	AV733574 AV733574
c	37	40	3.2	895	13	BQ729093	BQ729093 AGENCOURT
c	38	40	3.2	1101	29	CNS00ESI	AL069797 Drosophil
c	39	40	3.2	1412	28	BZ569852	BZ569852 msh2_1049
c	40	39.8	3.2	828	28	AZ686097	AZ686097 ENTLI59TF
c	41	39.6	3.2	780	10	BE642964	BE642964 Cri2_7I0
c	42	39.6	3.2	1024	28	CC437374	CC437374 PUHNW25TD
c	43	39.4	3.2	387	14	CA859923	CA859923 EST634417
c	44	39.4	3.2	646	9	AL045329	AL045329 DKFZp434A
c	45	39.4	3.2	715	28	BH924586	BH924586 odi55e06.

ALIGNMENTS

RESULT 1
BX416727/c

```
LOCUS       BX416727                712 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  CS0DA011Y14 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
            CS0DA011Y14 5-PRIME, mRNA sequence.
ACCESSION   BX416727
VERSION     BX416727.1  GI:3765629
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 712)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fullength.invitrogen.com/ InvitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DA011BE07QP1.
FEATURES             Location/Qualifiers
     source          1..712
                     /organism="Homo sapiens"
```

```
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DA011Y114"
                /tissue_type="NEUROBLASTOMA"
                /clone_lib="Homo sapiens NEUROBLASTOMA"
                /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                with a NotI-oligo(dT) primer. Five prime end enriched,
                double-strand cDNA was digested with Not I and cloned into
                the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                Library was not normalized."

ORIGIN

Query Match      4.1%;  Score 51.6;  DB 13;  Length 712;
Best Local Similarity  12.1%;  Pred. No. 0.06;
Matches   49;  Conservative 180;  Mismatches 175;  Indels   0;  Gaps   0;

Qy    432  ACCATATAAAGAAGAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTGGAATATACTGT  491
Db    709  MMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM  650

Qy    492  ACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAAGATACTAAGCT  551
Db    649  MMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM  590

Qy    552  ATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTAGCTCAAGCACA  611
Db    589  MMMMMMMMMMMMMMMMMMMTMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM  530

Qy    612  AAGCATTTTAAACAAACAAACCCACCCAGGCTATACGATTTATGAACGTGACTCCTCAATCGT  671
Db    529  MHMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMCMCMWMMMMM  470

Qy    672  CACTCATGACAATGACATTTTCCGTACGATTTTACCAATGACAAGAGTTTACTTACCA  731
Db    469  MMAMAMAMAMMMMMMAMAMAMMMMMMMMMAMAMMMMMAAAMAAAMAAAAAAAAAAAAA  410

Qy    732  TGTCAAAAATCGGGACAAGCTTATGAGATCAAGAGTTTACTTACTACCA  791
Db    409  AMAMAMAAAAAAAAAAAAAWAAAAAAAAAAAAAWAAAAAAWAAAAAAAA  350

Qy    792  AAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAA  835
Db    349  AAAAAAAAAAAAAAWWAATWWAAWWWWTWWWTWWTWWTWW  306

RESULT 2
BX436282/c
LOCUS       BX436282                1124 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  BX436282 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP001YC01
            5-PRIME, mRNA sequence.
ACCESSION   BX436282
VERSION     BX436282.1  GI:30787521
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1124)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0CAP001AB01QP1.
FEATURES             Location/Qualifiers
     source          1..1124
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

```
                /db_xref="taxon:9606"
                /clone="CS0CAP001YC01"
                /tissue_type="THYMUS"
                /clone_lib="Homo sapiens THYMUS"
                /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                with a NotI-oligo(dT) primer. Five prime end enriched,
                double-strand cDNA was digested with Not I and cloned into
                the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                Library was not normalized."

ORIGIN

Query Match      4.0%;  Score 49.4;  DB 13;  Length 1124;
Best Local Similarity  22.2%;  Pred. No. 0.21;
Matches  129;  Conservative 174;  Mismatches 278;  Indels   0;  Gaps   0;

Qy    492  ACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAAGATACTAAGCT  551
Db   1008  AYYYYYYYYYAAAAAAAAAAAAAAHMAAAYYYMYHHYYYYYAAAAAAAAAAAYYY  949

Qy    552  ATTGAAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTAGCTCAAGCACA  611
Db    948  YYYHAAAYYYYYYYYYYYYYYYYYAYYYHAAAAHHHHHAAAAAAAHAAAAHAAAAA  889

Qy    612  AAGCATTTTAAACAAACAAACCCACCCAGGCTATACGATTTATGAACGTGACTCCTCAATCGT  671
Db    888  AHYYYYYYYYYHHHHAAAAYYYYYYHHAAAAAAYYYYYYYYY  829

Qy    672  CACTCATGACAATGACATTTTCCGTACGATTTTACCAATGACAAGAGTTTACTTACCA  731
Db    828  YYAAAAAAAAAYYYYYYYYYYYAAAAAAAAHAAAAYYYARAAYYYYYYYAAAY  769

Qy    732  TGTCAAAAATCGGGACAAGCTTATGAGATCAATAAAAATCTGGTCTGAATGAAGAAAT  791
Db    768  YYYYAAAMHYHAAAAAAAHYYYYAAAAAAHYYYYAAAAAAAAAHHAAAAAAAAAAAA  709

Qy    792  AAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGGGAAAAGCCGTA  851
Db    708  AAAHYYYYYYYYYYYYYYYYYYYYYYYHAAAAAAAAAAAHHAAAAAAAAAAA  649

Qy    852  TGATCCCTTGATCGCCAGTCACTTGAAAACTGTTCACCCATCAAATACGTTGATGTGCAACAC  911
Db    648  AAHYYYYYYYYYYYYYYYHAHHAAAAAAAAYYYYYAHYYYYYYYAAAAAAH  589

Qy    912  CAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGTAACTTAGACTTCAG  971
Db    588  YHAMYYYYYYYYYYHAAAAAAAAAAAAAAAHHYYYYYYHAAA  529

Qy    972  AGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACACAATCCGATGCTTTTGG  1031
Db    528  AAAAAAAAAAAAAAAAAAHAAAAYYAAAHYYHAAHHAAAAYYYYYYYYY  469

Qy   1032  TATTATGGACTATACCTTAACTGGGAAAAGTAGAGATAATC  1072
Db    468  YAAAAMAAAAAAAAAYYHAAAAAAHHAAAAAAAAAAAAAAYY  428

RESULT 3
BQ678748
LOCUS       BQ678748                1217 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  AGENCOURT_8241036 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6265192
            5', mRNA sequence.
ACCESSION   BQ678748
VERSION     BQ678748.1  GI:21791427
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1217)
  AUTHORS   NIH-MGC http://mgc.ncl.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

```
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2431 row: o  column: 17
High quality sequence stop: 143.
```

FEATURES   Location/Qualifiers
source   1..1217
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6265192"
            /tissue_type="melanotic melanoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_112"
            /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G). Library constructed by Ling Hong in the
            laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            Superscript II RT (Life Technologies). Note: this is a
            NIH_MGC Library."

ORIGIN

Query Match   4.0%; Score 49.4; DB 13; Length 1217;
Best Local Similarity 43.7%; Pred. No. 0.22;
Matches 218; Conservative 0; Mismatches 281; Indels 0; Gaps 0;

Qy   447 ACCAATACAAAATCAAGCGAAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTT 506
          |||| |||| |||| || ||| |||||| |||||| || |||
Db   282 AACACGAACAAACACAAAGGCAAACAAAGAAAACACACAAAAAAAAACCCCACCGA 341

Qy   507 AAACCCTGATGACGATTTCAGACCCAGTCTCAAAGATACTAAGCTATTGAAAACACTAGC 566
          |||| |||||
Db   342 AAAGGAAAGAAAAATAACTAACCCAGCAAGAAAGACAAAAATAACAAAATAAAAAAAAAA 401

Qy   567 TATCGGTGACACCATCACATCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAA 626
          ||||| ||||
Db   402 CAAAGGAAAACACCGCCACAAAAGAAAACCAACCAACCAAAGAACAAACCAATAATAACAC 461

Qy   627 AACCCACCCAGGCTATACGATTTATGAACGTGACTCCTCAATCGTCACTCATGACAATGA 686
          ||||| |||||
Db   462 AACAAACGAACACGAAACAGCAGAGAGAAACCGACAAATAACCACCATAACAAACAAGA 521

Qy   687 CATTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGA 746
          |||||| |||||
Db   522 AAAAAAAAAATACGAAAATATAGAGAAAAATAGAGAAGAAGAAATAAAAAAAAATAAGGA 581

Qy   747 ACAAGCTTATGAGATCAATAAAAATCTGGTCTGAATGAAGAAATAAACACTGACCT 806
          |||||| |||
Db   582 AAAAGAACAAAGAACACAGAAAAATACAGAAGAAAAATAGAAGAAAAAAACAAAAAAACACA 641

Qy   807 GATCTCTGAGAAATATTACGTCCTTAAAAAAGGGGAAAAGCCGTATGACCCTTTGATCG 866
          ||||| |||
Db   642 ACACCCAAACAAAACGACATGATGAAATAAACGAAATGCAAAATGCAATCAATATAACAC 701

Qy   867 CAGTCACTTGAAACTGTTCACCATCACAAATACGTTGATGTCAACACCCAACGAATTGCTAA 926
          ||||| ||||
Db   702 AAAACAACAAAAAACACAACCAACAACAGAAAACGAAAGATAAACAACTAACACAAATACACAA 761

Qy   927 AAGCGAGCAGCTCTTAACA 945
          |||| ||||
Db   762 TCAACAACAGCACAAAATA 780
```

```
RESULT 4
BX423299/c
LOCUS    BX423299    1175 bp    mRNA    linear    EST 13-MAY-2003
```

DEFINITION   BX423299 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
             CS0DF008YB06 3-PRIME, mRNA sequence.
ACCESSION   BX423299
VERSION   BX423299.1 GI:30655486
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1175)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. This sequence belongs to sequence cluster 8830.r For
             more information about this cluster, see
             http://www.genoscope.cns.fr/
             cgi-bin/cluster.cgi?seq=CS1AF002ZH04NP1&cluster=8830.r. Contact :
             Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS1AF002ZH04NP1.
FEATURES   Location/Qualifiers
source   1..1175
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="CS0DF008YB06"
             /tissue_type="FETAL BRAIN"
             /dev_stage="fetal"
             /clone_lib="Homo sapiens FETAL BRAIN"
             /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
             was primed with a NotI-oligo(dT) primer. Five prime end
             enriched, double-strand cDNA was digested with Not I and
             cloned into the Not I and EcoRV sites of the pCMVSPORT 6
             vector. Library was not normalized."

ORIGIN

Query Match   3.9%; Score 48; DB 13; Length 1175;
Best Local Similarity 36.2%; Pred. No. 0.48;
Matches 117; Conservative 58; Mismatches 145; Indels 3; Gaps 1;

```
Qy    581 TCACATCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCACCAGGCT 640
          : ||:|| ||||:||||::|                       ::||||:  ::|||||
Db   1098 TANYTTTCMAAAATWAAAAAYTTAWAWATAARCAAABTAAAAMAARCCCVWAWWCAA 1039

Qy    641 ATACGATTTATGAACGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGA 700
          ||:|||||:| |||:||||                        :|::||::||:|||:
Db   1038 YATTATTTAWTBMAAWHAWATATYATWTAAMAAWTATAAAAAWTAAWAATTAHTA 979

Qy    701 TTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGA 760
          :||||:                                      ||||: ::|||||::
Db    978 WTTTWAAWAAAATTAAATWTTTTTWAATYTWTYMAWTAAAWYWTWAAAAAAWTAAATT 919

Qy    761 TCAATAAAAATCTGGTCTGAATGAAGAAATAAACACTGACCTGATCTCTGAGAAAT 820
          :|||::|||| ||:|||:                       :||||||||||||:
Db    918 TYAAWAAWHTWAAAA---TAATTTWAAAAATAATAATYAAAYAAAAAMWYYCAAAATY 862

Qy    821 ATTACGTCCTTAAAAAAGGGGAAAAGCCGTATGACCCTTTGATCGGCAGTCACTTGAAAC 880
          ||||:||                                     ::|||||::|||||
Db    861 AAMHVATCCATATHCAATCAATTYYTYATYCCATCYCTTYAKCBWMBTCTTTMATWA 802

Qy    881 TGTTCACCATCAAATACGTTGAT 903
          ||||||                ||
Db    801 TTAWWAAAATTTWTAAKYTTCWT 779
```

```
RESULT 5
CNS0100X/c
LOCUS    CNS0100X    1101 bp    DNA    linear    GSS 26-JUL-1999
```

```
DEFINITION   Drosophila melanogaster genome survey sequence SP6 end of BAC
             BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit
             fly), genomic survey sequence.
ACCESSION    AL098379
VERSION      AL098379.1   GI:5609990
KEYWORDS     GSS.
SOURCE       Drosophila melanogaster (fruit fly)
ORGANISM     Drosophila melanogaster
             Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
             Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
             Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1  (bases 1 to 1101)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
             collaboration with the European Drosophila Genome Project (EDGP) -
             http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
             library (Dros BAC) was made by Alain Billaud at CEPH (Centre
             d'Etude du Polymorphisme Humain) with funding provided by a MRC
             project grant. The DNA was prepared from embryos by Alain Bucheton
             and Genevieve Payan. It has been constructed in the vector
             pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN03G04"
                     /clone_lib="DrosBAC"
                     /plasmId="pBeloBAC11"
                     /note="end : SP6"
ORIGIN

Query Match        3.8%;  Score 47.8;  DB 29;  Length 1101;
Best Local Similarity  16.9%;  Pred. No. 0.53;
Matches     68;  Conservative 172;  Mismatches 163;  Indels 0;  Gaps 0;

Qy    435 ATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTGGAATATACTGTACA 494
              :::::::::::::::: :::::::::::::::::::  ::  ::: :::::
Db    833 ACAHTWWWTAMDEMMTMMMCTYHMWCMYAHHHYHYCTMWEIMHTHIWHCHMGHCHYMT 774

Qy    495 GTTACTCCCTTAAACCCTGATGACGAATTTCAGACCAGGTCTCAAAGATACTAAGCTATT 554
              :::::::::: ::  ::  :::  ::::: ::: :::: :  ::: :::::
Db    773 MHHHHYHTMYTAAMMAAAAAAAAAAMAAAAAAAMAAMATWTATMWTTATTTATHTTT 714

Qy    555 GAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTAGCTCAAGCACAAAG 614
              :::::::: : ::   :::: :::: :::: :: ::: ::::: :: ::::
Db    713 TYATMTWYWTWTYHCMTWSHCHCCACMHWYHCHCYMCHYMWHAWMHWCWMAHACAMAMHA 654

Qy    615 CATTTTAAACAAAACCACCACCAGGCTATACGATTTATGAACGTGACTCCTCAATCGTCAC 674
              ::    ::::::::::::::::::::  ::::: :  :: :: ::::: :  ::
Db    653 MMMWAMAAMMWCCMMCWCWCYMYCTMAAHTMWHWMHWHMWHWHHAAHTMRAAACWCHAM 594

Qy    675 TCATGACAATGACATTTTCCGTACGAGATTTTACCCATGAGTTTACCTACCATGT 734
              ::  : ::: :::::::::::::  :::  ::::::: :  :::: :::: ::::
Db    593 ATHAHTMCHTHMHMTNTYTHTYTCHAYHTYWHCWCWAHAMYKCYHAMMMACAYAMAMMAH 534

Qy    735 CAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAA 794
              :::  ::::::::::::: ::: ::::::: :::: ::: :::
Db    533 AMAYMWMMTMAMHMHMIAWWCTNHTWBHAWMHAHMHAAMWMHMWAHAWAACHHHMWMAMAHW 474

Qy    795 CAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAA 837
              :::::  :: :: ::::::::::: ::::::::::::::
Db    473 MMWCAAAAWAWAWAAMWAAMWAHAWWHAHTMCWAHAHAWMWAHHAYAH 431

Qy   1001 AACTTACTCTACACAATCTCGATGCTTTCGTATTATGGACTAATGGACTGAAAAAG 1060
              :::::::: :::::::::::::::::     ::::::::::::::::::::::::
Db    857 HHTMWATMHMYAWWTAHWAHWATTTWYTTWTWHTTWWWTATHWHWHHCWHTHTHMMM 916

Qy   1061 TAGAGGATAATCACGGATGACCATGACCACCGTATCATAACCGTTT 1102
```

```
ACCESSION    BX414650
VERSION      BX414650.1   GI:30763455
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 994)
AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. This sequence belongs to sequence cluster 6015.f
             Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS0CAP001DG01NP1.
FEATURES             Location/Qualifiers
     source          1..994
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0CAP001YN02"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

Query Match        3.8%;  Score 47;  DB 13;  Length 994;
Best Local Similarity  17.8%;  Pred. No. 0.83;
Matches     93;  Conservative 191;  Mismatches 238;  Indels 0;  Gaps 0;

Qy    581 TCACATCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCT 640
              :::: :: :::::::::::::::::::::::::::::::::::::::::::::
Db    437 YCACMCCCAAHMHAHAAAAAAAAACAMAYCYMYATWAWTWMAMANNYMNYHYYNNTTYYT 496

Qy    641 ATACGATTTATGAACGTGACTCCTCAATCGTCACTGCATGACAATGACATTTTCCGTACGA 700
              :: : :: :  :::::::::::::::::::::::::::::::::::::::::::::::::
Db    497 TTTYYYYYYYYYYYCCYYYYYYYYYYYYYTYTTHWAHYA 556

Qy    701 TTTTACCAATGGATCAAGAGTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGA 760
              :::::::::::::::::::::::: ::::: :::: ::::::::
Db    557 AAAAMMWAMWAMMAMMMAATTTWAAAAAAAAAAAAWHHTWHAWWTAAAAMWAMAAA 616

Qy    761 TCAATAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAAT 820
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db    617 AAAAAAAAAAAAAAATAMWMMWTMHTTWTAAAAAAAAAWTTWWHHYTHHMYCMCMTAHWTT 676

Qy    821 ATTACGTCCTTAAAAAAGGGGAAAAGCCGTATGAGTCCCTTGATCGCAGTCACTCGAAAC 880
              ::::::::::::::::::::::: :::::  :::::::::::::::::::::::::::::::
Db    677 TTTWEMHMWMWWATWWTATTTTTYTTTHTHMYCCTCHWWTTYTMEMC 736

Qy    881 TGTTCCACCATCAAAATACGTTGATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCT 940
              :::::::::::::::::: ::::::::::::::::::::::::::::::::::::::::::::
Db    737 TTHTYCCYCWHWWTTTTTTYWHTYWTYMHWAHWTTWTWHTTWTHWMHWTTWMW 796

Qy    941 TAACAGCTAGCGACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAAGGCTA 1000
              ::::::: ::::::::: ::::::::::::::::::::::::::::::::::::::::::::
Db    797 TTTTYWWTYTTWWYTTWWYTTWHHTTWTTWTTHWATHTWWWTAWCTWHAWTMWWWAAAAAVMTHHHT 856
```

```
RESULT 6
BX414650
LOCUS       BX414650    994 bp    mRNA    linear    EST 15-MAY-2003
DEFINITION  BX414650 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP001YN02
```

```
Db  917 AAMAAAAWHTWMYMYMTTWMTTWWTAWYHWYWTTTWHHTWHTYYYYT 958
            |        ::  :: ::      ::::::     |:::|::|::|
```

RESULT 7
CNS0039G

```
LOCUS       CNS0039G                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION     AL063921.1  GI:4941778
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp; the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR08K10"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN
```

```
Query Match           3.7%;  Score 46.4;  DB 29;  Length 1101;
Best Local Similarity 16.3%;  Pred. No. 1.2;
Matches  98;  Conservative 240;  Mismatches 264;  Indels 0;  Gaps 0;

Qy  436 TATAAGAGAAAAACCAATACAAAATCAAGCGAAAATCTGTTGATGTGGAATATACTGTACAG 495
             | |  ||||: ::: |:::  |||||                |||
Db  446 WWAAAWWAAAAAWTAWWTTAWTWAAAAAAAAAAAATTWTTTTTTWTTTWTTTAWTTWATAW 505

Qy  496 TTTACTCCCTTAAACCCTGATGACGATTTCAGACGACGTCTCAAAGATACTAAGCTATTG 555
         :||::  ||                  :  |::                  ::
Db  506 WTTWWWTTAAAWAAAAAAAAAAAAAWAAAWAAAWATAAAWTWWTWWTTYTTWWAWATAAAM 565

Qy  556 AAAACACTAGCTATCGGTGACACCATCACCATCCAAGAATTACTAGCTCAAGCACAAAGC 615
         :||:|   ||     ||| : |||||      :|:|:  |::: | ::: :|:|: :
Db  566 CMAAWYHYTYTYTYHYTYWTYTMTWYHTMTWYHTMYTHAWAHTTWYHWYHTYAMWHWMTWHWT 625

Qy  616 ATTTTAAACAAAAACCCCACCCAGGCTATACGATTTATGAACGTGACTCCTCAATCGTCACT 675
         :: ::: || ||||:  ||:  || :::: ::       ||:  ||:|::   | ::
Db  626 MWAWWHWTTYTRAYYYYYTCMYYTHYHYMWHHAHAHAAAWTTHTWWTHAYHWATYHYYY 685

Qy  676 CATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCATGTC 735
         : ||||:| |:|:|| :|:   |||:|:|| ::     ||::| |||||||:::: : :
Db  686 MYCAMMCMCTHTCHHCYYYHHYTAHHTHTHHWYAHYMWYMWAAYWMYCTACTYHYHH 745
```

```
Qy  736 AAAAATCGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTGAATGAAGAATAAAC 795
         :::: :  :: |: ::      ::::::     :::::||| | ::::::||::||
Db  746 HHHYHWAYHTTWYAWAHAMWMWWHHAHYAAAAAWAAWATTHHYHHTHTHYMHHTYMYHYMY 805

Qy  796 AACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGGGAAAAGCCGTATGAT 855
         ::  :  :::|| ::: ::  |: |:|: :::::: |:: ::||   |::::::
Db  806 TCCYMCTYHCWHYHTAYTCWTWTHHWMWTWTWTHTWTTHWTTHWTTTCWHTHTTTCWCWWW 865

Qy  856 CCCTTTGATCGGCAGTCACTTGAAACTGTTCACCATCAAATACGTTGATGTCAACACCAAC 915
         : :|:||| :|  | ::  ::|: | :|::  ||:::|| :: |:||::|| || :::
Db  866 HATTWTTWATHCWACMTWHWHWMWMHHHHHACHAHHTHMCMCHHHMHCTCHHHHTMYHM 925

Qy  916 GAATTGCTAAAAAGGCAGCAGCTCTTAACAGCGTAGCGAACGTAACTTAGACTTCAGAGAT 975
         || :||  :|:|: ::|: ||::::|| :|:||     :|::||:|: |:|:||:||: |:
Db  926 TCHWMWHMWHHWHWWMAATWMTTMTTWMAMCCMMHHCHMYHMHMYMYMYCCHYYCTCHTH 985

Qy  976 TTATACGATCCTCGTGATAAGGCTAAACTACTCTCACAACAATCTCGATGCTTTTGGTATT 1035
         :: |: :| | ||: |: ::|: |:| |: |::: |:|| |:|: |:::| ::| ||:::
Db  986 ATTHYHYMCTCTHYCTWHTYHYCTWMHTYTAYWWAWTAHAMTTATWWWMEWWAHWATWWWWWWAT 1047

Qy 1036 AT 1037
         |:
Db 1046 AW 1047
```

RESULT 8
BX425603/c

```
LOCUS       BX425603                 885 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  BX425603 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
            CL0BB022ZA07 3-PRIME, mRNA sequence.
ACCESSION   BX425603
VERSION     BX425603.1  GI:30770486
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 885)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CL0BB022ZA07FP1.
FEATURES             Location/Qualifiers
     source          1..885
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CL0BB022ZA07"
                     /clone_lib="Homo sapiens NEUROBLASTOMA"
                     /tissue_type="NEUROBLASTOMA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN
```

```
Query Match           3.7%;  Score 46.2;  DB 13;  Length 885;
Best Local Similarity 7.3%;  Pred. No. 1.3;
Matches  45;  Conservative 162;  Mismatches 412;  Indels 0;  Gaps 0;

Qy  181 AAACCATTTGCTACTACTGATAGTGGCGGCGATGCCACCATARACTTGAAAAAGCTGACTACTA 240
         :::::
Db  804 MMMMMNMNNMNNNNNNNNNNNNNNNNNNNNNNNNNNNNMNNMNMMMMMMMMMMMMMM 745

Qy  241 AAGGCTATTCAAGAACAATTGATCGCTAACGTCCCACAGTAACGACGACTACTTTGAGGTC 300
```

with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

```
Query Match     3.6%;  Score 45.4;  DB 13;  Length 1085;
Best Local Similarity  31.2%;  Pred. No. 2.1;
Matches  104;  Conservative  69;  Mismatches 160;  Indels 0;  Gaps 0;

Qy   439 AAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTGGAATATACTGTACAGTTT 498
             |||||||  ||| :|:    ::|:  |||
Db  1066 AAAYAMAAACWAAAMMAAAYCCCHWCWMWAAATWHYTAWAMYWCWAWTTHATTCTWATMTY 1007

Qy   499 ACTCCCTAAACCCTGATGACGATTTCAGACCAGGTCTCAAAGATACTAAGCTATTCGAAA 558
          ::  |:  | :|:|||:|||:|| |||||| :: || |||  ||||  ||
Db  1006 MYACYAAAAMCMCYCHAMACCYMYCYHAHAMCYCYTYAAAAAMCAAAAAATAWAMAAAAMH 947

Qy   559 ACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTAGCTCAAGCACAAAGCATT 618
          :|:  || ||::   :|:  || |  :|:|  | |||||:|||:||
Db   946 MCAWAATACCCMCTCCTAHAATAWAAAAATAAMTTATYWMMYTCYMYCYCWCTCTAACYC 887

Qy   619 TTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGACTCCTCAATGTCACTCAT 678
          :|:  || |:  || :|:|   |:|| ::|| || |||
Db   886 ATAATTAAAAHAAAYWHHTAMAHATTHTCTMTWWAAAAMWWAYATAACCCTTHTTAAT 827

Qy   679 GACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAGTTTTACTTACCCATGTCAAA 738
          :|:  ||:|:   |:|||:|| :||:|:|||||| |||||  ||
Db   826 YTWAAAYAMCTTAAAATTYTAAAYCTYYAAMTCTCCMAAAYAATAAAAMCCACHACAMY 767

Qy   739 AATCGGGAACAAGCTTATGAGATCAATAAAAAA 771
          |||| :  ::|:| | |:  |:| |
Db   766 ATAYYAMAAVAAYYCYYAMTCCMWAAAACMMA 734
```

RESULT 10
BX361080/c
LOCUS   BX361080   1201 bp   mRNA   linear   EST 05-MAY-2003
DEFINITION BX361080 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI079YN16 3-PRIME, mRNA sequence.
ACCESSION   BX361080
VERSION   BX361080.1 GI:30374504
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI079DG08NP1.
FEATURES   Location/Qualifiers
     source   1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI079YN16"
                     /tissue_type="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

```
Db   744 MMMMMNNNNMMMMMNNNNKGNKGNKKGKMMNNNNNKKGNNNNNNNNNNNKMMNNNKK 685
              ::::    |:     ::   :
Qy   301 ATTGATTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTTGCTGAC 360
            ::  ::::::  ||               ::      ::    :
Db   684 NKGKNKKKKGKKGKKGKKQNTNKMNNNNNNNNMMMMMMMMMMMMMMMM 625

Qy   361 AAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCCGACAT 420
          |:|:  :                                      ::::::
Db   624 NAMMMNAMMMAMMMMMMMMMMMMMNNNNNNNNNNNNNNNNNKKKKKKNKNNNN 565

Qy   421 GTGCGCGTTAGACCTATATAAAGAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
              ::  :::::::   ::::::  ::::    :: :::
Db   564 NNNNNMMNNNNNNNNNMMMMMAMMMNNNNNNNNN 505

Qy   481 GAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAA 540
          ::  ::: :::|::|   :::: ::::   :::: ::
Db   504 NNNNNNNNNNKNNNNKNNNNMMMMMMMMMMNCMMMMMNNMMMMMMMMMMMNNNN 445

Qy   541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGGTGACACCATCACATCTCAAGAATTACTA 600
          :|:  :      :::::::::
Db   444 NNMNMMCMKMNNNMKNKNNNNNNNNNNNTKNTNNNNNNNNNNNNNNNNNNNNT 385

Qy   601 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC 660
           :::    :::|:   :::::|:  :: :::
Db   384 TNNNNMNNNNMNNNNNMMAMMMMMAMCMMAMMMAAANMMMNANMMMMMMMMMM 325

Qy   661 TCCTCAATGTCACTCATGACAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 720
                                ||                  ||||||
Db   324 NNNNNNNNNCSNNNGNNNKNGGNNNCNCNCCCCCCNTNNNNNNNNNNCNNCNNTN 265

Qy   721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAAATCTGGTCTG 780
              ::::::    ||  ||||     :::   ||||||
Db   264 NTNNNAAAAAAAMMMMMWWAAAAAAAAAAAAMNNNNNNANAANAAAAAAAAAAANAA 205

Qy   781 AATGAAGAAATAAACAACA 799
          ||  ||||||    |||
Db   204 AANAAAAAAAAAAAAAAAAAA 186
```

RESULT 9
BX414831/c
LOCUS   BX414831   1085 bp   mRNA   linear   EST 15-MAY-2003
DEFINITION BX414831 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP002YL02
            5-PRIME, mRNA sequence.
ACCESSION   BX414831
VERSION   BX414831.1 GI:30767501
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1085)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0CAP002DF01QP1.
FEATURES   Location/Qualifiers
     source   1..1085
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0CAP002YL02"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="vector: pCMVSPORT_6; 1st strand cDNA was primed

Query Match
Best Local Similarity   3.6%;   Score 45.4; DB 13; Length 1201;
                       12.4%;   Pred. No. 2.1;
Matches 27; Conservative 111; Mismatches 80; Indels 0; Gaps 0;

```
Qy  427 GTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTGGAATAT 486
            :::|:| ||| ||||||||||:::::::: :::: :: ::::: :
Db  976 KKKAKAADAAAAAAWAAARAAAMWAMAMMMMMMNKMKKKKHMMMMKKMKMMMDKKKKWMM 917

Qy  487 ACTGTACAGTTTACTCCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAAGATACT 546
             :::::: ::: ::::::  :: :||:|:: :: ::: :::|:|
Db  916 MMMMMMMMT-TTHMMMMMTWMMMTKTMMKTMWMMKTMWMMKMMHMKMMMHM 857

Qy  547 AAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATTCAAGAATTACTAGCTCAA 606
             :: :::  ||:|  :: :::  :: ::|:: :::|::
Db  856 MMMMMTMKKMMHNMMMMHNKMKKKKKKCHHKKMTNKKMMMMMMMMMYMMKMTTMMM 797

Qy  607 GCACAAAGCATTTTAAACAAAACCCACCCAGGCTATAC 644
             :: ||  :::: ::|  :::
Db  796 KKBTKYMKKKTTCMMMMTMKMMCTTTCCCCCTTYMC 759
```

RESULT 11
CNS06X9S
LOCUS   CNS06X9S   1007 bp   DNA   linear   GSS 06-JUL-2001
DEFINITION T3 end of clone AX0AA039F08 of library AX0AA from strain CBS 7064
            of Pichia farinosa, genomic survey sequence.
ACCESSION   AL419462
VERSION   AL419462.1 GI:12202640
KEYWORDS   GSS.
SOURCE   Pichia farinosa
  ORGANISM Pichia farinosa
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE   1 (bases 1 to 1007)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE   Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2 (bases 1 to 1007)
  AUTHORS   de Montigny,J., Spehner,C., Souciet,J., Tekaia,F., Dujon,B.,
            Wincker,P., Artiguenave,F. and Potier,S.
  TITLE   Genomic exploration of the hemiascomycetous yeasts: 15. Pichia
            sorbitophila
  JOURNAL   FEBS Lett. 487 (1), 87-90 (2000)
  MEDLINE   20584725
   PUBMED   11152890
REFERENCE   3 (bases 1 to 1007)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT   This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES   Location/Qualifiers
     source   1..1007
                     /organism="Pichia farinosa"
                     /mol_type="genomic DNA"
                     /strain="CBS 7064"
                     /db_xref="taxon:4920"
                     /clone="AX0AA039F08"
                     /clone_lib="AX0AA"
                     /note="end : T3"
ORIGIN

Query Match
Best Local Similarity   3.6%;   Score 44.6; DB 29; Length 1007;
                       35.0%;   Pred. No. 3.2;
Matches 143; Conservative 51; Mismatches 215; Indels 0; Gaps 0;

```
Qy  437 ATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTGGAATATACTGTACAGT 496
            :|: |:|:::|| |||||| :: ::||||| ::|:|
Db  205 AAWTAAAMAWMWAWAAACAMGTRCATGTWCMWGGTGATGAWMGTRAWTATWAWGAWAAAR 264

Qy  497 TTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTCAAAGATACTAAGCTATTGA 556
            :| |||  |:| :||  |||
Db  265 AWARTAAWGGWTAAAGMWAWAWAAAAAWRGAAGWGGTTATAAWAAWAAWAAWAAWRAWA 324

Qy  557 AAACACTAGCTATCGGTGACACCATCACATTCAAGAATTACTAGCTCAAGCACAAAGCA 616
            :|: |||
Db  325 AWAAAAAAAAAAGCAMAAAAAWAAAAAAWAAAAAAWAAAAAAACGAAAWWAA 384

Qy  617 TTTTAAACAAAACCCAGGCTATACGATTTATGAACGTGACTCCTCAATCGTCACTC 676
            |:|| |||| :::
Db  385 TGTAAAAAAAAATAAAATAAAAWATAAAAAWWAWAAAAAAAAAWTAAAWAAAT 444

Qy  677 ATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCATGTCA 736
            : |: |||||| |||
Db  445 GAAATAAAAAGATTTAAWWAWAAAAWAWAWAWWAWAAAAAAAATAAAAAAAAAAAAAAWA 504

Qy  737 AAAATCGGGAACAAGCGTTATGAGATCATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACA 796
            |||| |:|: ||||
Db  505 AAAAAAAWATWTWAACWAMAAAAAAAAAAAGAWTAWAAAAAAATWWAWAAAAWWW 564

Qy  797 ACACTGACCTGATCCTCGAGAAATATTACGTCCTTAAAAAGGGGAAA 845
            ||| |||||
Db  565 AAAAAAWWATAAATACGAATGATTATTAAACAACGTTTAGTGGGGAATAA 613
```

RESULT 12
CNS00LT2/c
LOCUS   CNS00LT2   1101 bp   DNA   linear   GSS 14-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR49P19 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL078714
VERSION   AL078714.1 GI:5102004
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp. The same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or

FEATURES
   source
           filters for hybridization from the BACPAC Resource Center can be
           found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
           Location/Qualifiers
           1. .1101
           /organism="Drosophila melanogaster"
           /mol_type="genomic DNA"
           /db_xref="taxon:7227"
           /clone="BACR48P19"
           /clone_lib="RPCI-98"
           /note="end : TET3"

ORIGIN

Query Match   3.6%;   Score 44.4; DB 29; Length 1101;
Best Local Similarity   25.7%; Pred. No. 3.6;
Matches 89; Conservative 96; Mismatches 161; Indels 0; Gaps 0;

Qy 551 TATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTACTAGTCAAGCAC 610
Db 836 TTTTHWMAMMAYHTTMWMMTTWMWMMMCMMMMAAAATTMMMMMAMHMTHHHCTMMMCMM 777

Qy 611 AAAGCATTTTAAACAAAACCACCACCAGGCTATACGATTTATGAACGTGACTCCTCAATCG 670
Db 776 CWCCMCMWMMMWMMMMCMMCMMCMCMMMCCMMYHMMHTTTWTTMMHMMAMATMMTCHMMM 717

Qy 671 TCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAAGAGTTTACTTACC 730
Db 716 THMWMHTTTMMMMWMMAAAWWTTWAHAAAWTTTTTTATTTAATAAAAAAAAMMTAA 657

Qy 731 ATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAAATCTGGTCTGAATGAAGAAA 790
Db 656 AAAAAAATTWAWAAAWAAAATTWAAAAAWAAAWAAAATAHWTWMWMWMTWAMMTAA 597

Qy 791 TAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGGGAAAAGCCGT 850
Db 596 WAMAAAAWATTAAAAAWWTTWTTAMTTWAWHMAMATMWAAWMAAWMAAAAAATATTAA 537

Qy 851 ATGATCCCTTTGATCGCAGTCACTTGAAAACTGTTCACCATCAAATA 896
Db 536 TWAAWAAWTAAAAAHWWTTWMHHTTHAAAAWAAAATMWAAAAAAAA 491

RESULT 13
CNS0187R
LOCUS   CNS0187R   878 bp   DNA   linear   GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
          BACN04B04 of DrosBAC library from Drosophila melanogaster (fruit
          fly), genomic survey sequence.
ACCESSION   AL108993
VERSION   AL108993.1 GI:5529297
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 878)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
          BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
          - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
          collaboration with the European Drosophila Genome Project (EDGP) -
          http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
          library (Dros BAC) was made by Alain Billaud at CEPH (Centre
          d'Etude du Polymorphisme Humain) with funding by a MRC
          project grant. The DNA was prepared from embryos by Alain Bucheton
          and Genevieve Payan. It has been constructed in the vector
          pBeloBAC11.
FEATURES
   source   Location/Qualifiers
          1. .878
          /organism="Drosophila melanogaster"
          /mol_type="genomic DNA"

          /db_xref="taxon:7227"
          /clone="BACN04B04"
          /clone_lib="DrosBAC"
          /plasmid="pBeloBAC11"
          /note="end : SP6"
ORIGIN

Query Match   3.6%;   Score 44.2; DB 29; Length 878;
Best Local Similarity   21.8%; Pred. No. 4;
Matches 73; Conservative 107; Mismatches 155; Indels 0; Gaps 0;

Qy 538 AAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTA 597
Db 43 ARAWMAATCAAAWRTATGAWAAAAAMMWMCAAAAAAAAAMMWMAAAAWAWAAAAMWMW 102

Qy 598 CTAGTCAAGCACAAAGCATTTTAAACAAAACCACCACCCAGGCTATACGATTTATGAACGT 657
Db 103 MRAAAAAAWWAMARGAWAWGCWGGAWWMWMAWRGMWMWWAWWAAAAAWWMWCTTTG 162

Qy 658 GACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTACCAATGGATCAA 717
Db 163 CMCTGAWTRTWGTTRTWTWAWWMRTTRAAWWMAAWWMAAWWWARGWWWWRGMWMAAA 222

Qy 718 GAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGT 777
Db 223 ARAWWWAMMWRCRCWTWCRCMAAMMAAMMAAMAAAAMAAMCAWMMWMWMAAMWCRAAW 282

Qy 778 CTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAA 837
Db 283 MGGAWARAAWTRGAWWMWRAMMAAAAMRAARMAWRAWMAAAWMAAATWTACAWAAAWAAA 342

Qy 838 GGGGAAAAGCCGTATGATCCCTTTGATCGCGAGTCA 872
Db 343 ARTMRTAWGMAGAAATAWGATATATATWACRCRCA 377

RESULT 14
CNS00HE9/c
LOCUS   CNS00HE9   1101 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC:
          BACR34A13 of RPCI-98 library from Drosophila melanogaster (fruit
          fly), genomic survey sequence.
ACCESSION   AL073374
VERSION   AL073374.1 GI:4953154
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
          BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
          - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
          collaboration with the Berkeley Drosophila Genome Project (BDGP).
          The BDGP is constructing a physical map of the Drosophila
          melanogaster genome using these BACs. For further information
          please see http://www.fruitfly.org The BDGP Drosophila
          melanogaster BAC library was prepared by Kazutoyo Osoegawa and
          Aaron Mammoser in Pieter de Jong's laboratory in the Department of
          Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
          NY. The library is named RPCI-98 and was constructed by partial
          EcoRI digestion of Drosophila DNA provided by the BDGP from the
          isogenic strain y2; cn bw sp, the same strain used for the BDGP's
          P1 and EST libraries. A more detailed description of the library
          and how to order individual BAC clones, the entire library, or
          filters for hybridization from the BACPAC Resource Center can be
          found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
   source   Location/Qualifiers
          1. .1101

us-09-882-509-3.rst

Page 9

```
                    /organism="Drosophila melanogaster"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:7227"
                    /clone="BACR34A13"
                    /clone_lib="RPCI-98"
                    /note="end : TET3"
ORIGIN

Query Match          3.6%;   Score 44.2;   DB 29;   Length 1101;
Best Local Similarity  32.7%;   Pred. No. 4.1;
Matches  109;   Conservative   7;   Mismatches  217;   Indels   0;   Gaps   0;

Qy    513  TGATGACGATTTCAGACCCAGGTCTCAAAGATACTAAGCTATTGAAAACACTAGCTATTCGG  572
                |:|| |:||| :|   | ::| |:   ||: || || |||:|
Db    428  TTAKGAWATTGYGARAAAATGGTGYGKARAATMTAAAANAAATAPARAAAAMNAAAAAWAAC  369

Qy    573  TGACACCACCATCACATCTCAAGATACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCA  632
                |||||||||||||  |||||  |||| ||||  ||||  |||||:|
Db    368  ACAAANNANNAAAAANANANAAAAANAANAAANNNCAANNNNNNAAAAAAAAAAAN  309

Qy    633  CCCAGGCTATACGATTTATGAACGTGACTCCTCAATCGTCACTCATGACAATGACATTTT  692
                ||||   |:|:||   |  ||  |::|::|:
Db    308  CGCNNNNNANANANANANNNNANANNNNNAAGNTNAAANNNNANNNNGNAANNNNNNAAANATG  249

Qy    693  CCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAAATCGGGAACAAGC  752
                ||:| |||  :||||  |:|||| :|:| |||
Db    248  NNNAAANNNNNGAAAAAANNAAANNAAAAANAAAAANAAANNAAAAANAAAAAA  189

Qy    753  TTATGAGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTC  812
                |||  :| |||   || |||:
Db    188  NAANANAAAAAANAANAAAAMAAAAANNCCAAAATNNNANANTAACANACNNCTNACAAA  129

Qy    813  TGAGAAATATTACGTCCTTAAAAAAGGGGAAAA  845
                |||||  |||||
Db    128  AGNNNNGAANNNNNANGNNANNAAAAGAGGAGNANA  96


RESULT 15
BX375784
LOCUS       BX375784               1109 bp     mRNA       linear   EST 08-MAY-2003
DEFINITION  BX375784 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC017YF05 5-PRIME, mRNA sequence.
ACCESSION   BX375784
VERSION     BX375784.1  GI:3044843
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1109)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 8559.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DC017CC03QP1&cluster=8559.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DC017CC03QP1.
FEATURES             Location/Qualifiers
     source          1..1109
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DC017YF05"
                     /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
```

```
                    /note="1st strand cDNA was primed with a NotI-oligo(dr)
                    primer. Five prime end enriched, double-strand cDNA was
                    digested with Not I and cloned into the Not I and EcoR V
                    sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match          3.6%;   Score 44.2;   DB 13;   Length 1109;
Best Local Similarity  34.8%;   Pred. No. 4.1;
Matches  103;   Conservative  31;   Mismatches  162;   Indels   0;   Gaps   0;

Qy    541  GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTACTA  600
                ||||||  || ||||||  ||  |
Db    625  GNTAAAAAANTGTTNNAAANNANCNAAATANANNNNAAANNNNNAATTTNAAAAAAAAAA  684

Qy    601  GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC  660
                |||||  |||  :||||  ||||
Db    685  GGTNAAAGGGTNAAAAATGNGNNAAAAANTKNNNANTCNNNTYNNTTTATNNNTCHAAT  744

Qy    661  TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG  720
                |::| ||:|:|  ::|   |:|   :|:|   |  |||::::  :
Db    745  TMYATTATYATHAAAATTAWAAATYCTCAYYTMYTAAATYYTTTMWTWAWWAACAAATC  804

Qy    721  TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG  780
                ||:|:  |||||| ::|  :|::||| ::|
Db    805  WTAATYCWAATYATYCYAAYCAAAAATCWTCAAWAAAAATCAMWATAATTCATCTC  864

Qy    781  AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAA  836
                  ::   ||  || |||  ||:|| |   ||||
Db    865  CYAYMTCCAAAAAHAAAAAYCCCTAACTMAAAAAAWTAAYAAATCTCTCAAAM  920


Search completed: May  6, 2004, 04:55:21
Job time : 3634 secs
```